

## SEQUENCE LISTING

### (1) GENERAL INFORMATION

(i) APPLICANT: JACOBSON, Myron K.; JACOBSON, Elaine L.; AMÉ, Jean-Christophe; LIN, Winston

(ii) TITLE OF INVENTION: GENES ENCODING SEVERAL POLY(ADP-RIBOSE) GLYCOHYDROLASE (PARG) ENZYMES, THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREACTIVE THEREWITH

(iii) NUMBER OF SEQUENCES: 38

#### (iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Fulbright & Jaworski L.L.P.  
(B) STREET: 666 Fifth Avenue  
(C) CITY: New York  
(D) STATE: New York  
(E) COUNTRY: USA  
(F) ZIP: 10103

#### (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
(B) COMPUTER: IBM PS/2  
(C) OPERATING SYSTEM: PC-DOS  
(D) SOFTWARE: Wordperfect

#### (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: To Be Assigned  
(B) FILING DATE: Concurrently Herewith

#### (viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: John E. Lynch  
(B) REGISTRATION NUMBER: 20,940  
(C) REFERENCE/DOCKET NUMBER: NIAD 201-JEL/ES

#### (ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 212-318-3000  
(B) TELEFAX: 212-752-5958

(2) INFORMATION FOR SEQ ID NO: 1:  
 (i) SEQUENCE CHARACTERISTICS:  
     (A) LENGTH: 4069 base pairs  
     (B) TYPE: nucleic acid  
     (C) STRANDEDNESS: single  
     (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```

1  accggaaggt gaacgaagcc cgaatcagaa cggtcatcc tgaggctggt aggggtgccg
61  tggaagaggg aaggcaggcg tctggatagg gcctggttcg ggaggctgtc agagcaggag
121 ctgcagaagc agtcagcggc agagggggca tggtgccggg aggcaccgag gaggggggcg
181 agtccgtccc tcccagggtt agtgaatgag gctctacgcc cgggctggcc cgggagactc
241 gtgtgcggg tcccagcatg agtgccggcc cgggtgtga gccctgcacc aagcgacccc
301 gctgggacgc cgctgcaact tctccgcggc ccgctcgga cgcocggagc ttcccggca
361 ggcagaggcg cgtcctcgat tccaaggacg ctccggtgca gttcagggtc ccgccgtcct
421 cgtcaggctg cgccctgggc cgggcgggac agcaccgagg cagcgccacc tctcttgttt
481 tcaaacagaa gactataacc agttggatgg acactaaagg aatcaagaca gttgaatcag
541 aaagtttgca tagtaaagaa aacaacaata caagagaaga atccatgatg agttctgtac
601 aaaaagataa cttttatcaa cataacatgg aaaaattaga aaatgtttct cagctagggt
661 ttgataagtc accagttgaa aaaggtacac agtatttgaa gcagcatcag actcggcta
721 tgtgtaagtg gcagaatgaa gggccacact cagaacggct tttggaaagt gaacctccag
781 cggtaactct ggtaccagag cagttcagta atgctaattg cgatcagtcg tccccaaagg
841 atgatcacag tgacacaaat agtgaggaga gtagagataa tcagcagttt ttgacacatg
901 taaagcttgc gaatgcaaag cagacgatgg aagatgaaca gggcagagaa gccagaagcc
961 accagaagtg tggcaaggct tgccatcctg cagaagcctg tgcagggtgt cagcaggagg
1021 agacagacgt ggtgtccgag agccccttgt cggacactgg ctctgaggat gttggtactg
1081 gactgaaaaa tgccaacaga ttgaatagac aagaaagtag tctaggaaat tctcctccat
1141 ttgagaaaga aagtgaacct gagtcaccaa tggatgtaga taattccaaa aatagttgtc
1201 aggattcaga agcagatgaa gagacaagtc caggttttga tgaacaggaa gatagcagtt
1261 ctgctcaaac agcaataaaa ccttcaagggt tccaaccaag agaagctgac actgagttga
1321 ggaagcggtc ctctgctaag ggaggtgaga ttcgattaca tttccaattt gaaggaggag
1381 agagtcgagc tggaatgaat gatgtgaatg ccaaacgacc tggaaagtact tctagcctga
1441 atgtagagtg cagaaattct aagcaacatg ggagaaaagg ttctaaaato acagatcatt
1501 tcatgagagt gcccaaagca gaggacaaaa gaaaagaaca atgtgaaatg aaacatcaaa
1561 gaacagaaaag gaagatccct aaatacattc cacctcacct ttctccagat aagaaatggc
1621 ttggaactcc tattgaggag atgaggagaa tgccaagggtg tgggatccgg ctgcctccct
1681 tgagaccatc tgccaatcac acagtgacta ttcgggtaga tcttttgcca ataggagaag
1741 ttctctaaac tttcccaaca cattttaaag atttgtgga caacaagcat gttaaagtgc
1801 cttgttcaga acaaaaactt taccctgtgg aagatgagaa tggtgagcga gctgcaggca
1861 gccggtggga actcattcag actgcacttc tcaacaggct cactcggccc cagaacctga
1921 aggatgctat tctgaagtac aatgtggcat attctaagaa atgggacttt acagctttga
1981 ttgatttctg ggataaggta ctagaagaag cagaagctca acacttgat cagtccatct
2041 tgcttgatat ggtgaaaatt gcactctgtc tgccaaatat ttgtacccag ccaataccac
2101 tcttgaaaca gaagatgaat cattccatca caatgtcaca ggaacagatt gccagctttt
2161 tagctaattg tttcttctgc acgtttccac gacgcaatgc caagatgaaa tcagagtatt
2221 ccagttatcc agatattaac ttcaatcggg tgtttgaagg acgttcatca aggaaaccag
2281 agaagcttaa aacgctcttc tgctaactta gaagagtcac agagaaaaaa cccactgggt
2341 tggtgacatt cacagacag agtcttgaag attttccaga gtgggaaaga tgtgaaaaac
2401 tctgactcgc actgcatgtc acttacgaag gtaccataga aggaaacggc cagggcacgc
2461 tacaggtgga ttttgcaaac cgtttcgttg gaggtggtgt aaccagtgcg ggacttgtgc
2521 aagaagaaat ccgcttttta atcaaccctg agttgattgt ttcacggctc ttactgagg
2581 tgctggatca caatgaatgt cttatcatca caggtactga gcagtacagt gaatacacag
2641 gctatgccga aacataccgc tgggcccggg gccatgaaga caggagcgaa agggacgact
2701 ggcagaggcg cagcactgag atcgtcgcca tcgacgccct ccacttcaga cgctacctcg
2761 accagtttgt gcccgagaag atcagacggg agcttaacaa ggcttactgt ggattttctt
2821 gtcttgaggt ttcttcagag aacctgtctg cagtggctac aggaaactgg ggctgtggtg
2881 cctttggggg tgatgctaga ctaaaagcct taatacagat cctggcagct gctgtagctg
2941 agcgagacgt ggtttatttc acctttgggg actcagaact gatgagagac atttacagca
3001 tgcatacatt cctcactgag aggaaactga ctgttgaga agtatataag ctgctgctac
3061 gatattacaa tgaagaatgc agaaaactgct ccacccccgg accagacatc aagctttatc

```

```

3121 cattcatata ccatgcagtt gagtcctgta cacagaccac caaccagccg ggacaaaagga
3181 cggggggcctg agggagccaag tgactagacg ctccccactt gtgtaacaag aaggtgtgac
3241 gtgtgaactg acatgatata catgtgtata taatccgcgt ttgtaggcaa ggatgcagtc
3301 ccttccgccc atgcagctgt cagtacatct gcgcctcctc catcccgact tacatagact
3361 gagacatact ttgtttcttt ttttttctat ttcagccctg attcttttat ttttctttct
3421 tttgcccata agacttcttg tgaaatttca tcagagtttg tgctcagcct ggcaggtgtc
3481 ttttttgatg cctaaatata caaatcacct ctgcagctag cagatgccac ggaaggtggt
3541 ggaaccctag gagctgtaac tgagtctgct gcagatctcc ctctgagcct ctcaccacct
3601 ccttattata attgtggttg tggaggtttt ttgatttttg aaataagagt tgggtttggt
3661 aaataataca gatctcctag gttaagagtt ttatatatta gaatactttt caaaaagtta
3721 ttttgagata tcacctttat ttgtaatggt aatttgccctg tcccttttcc cctgatcaat
3781 ttgtattgac tgttttttga aattgaccca aatgaaagga aatatgagaa taagagtttc
3841 ccaaattggtg tttaaaaaca aacaggttca agacacgcga aggacctcgt ttcctgggat
3901 ttttttttct tttctttttt tgaattagga ttattgtttg ttccttggtg cttgagacat
3961 attcatataa ccaaagttaa ggaactggga acttcgtggt gatttgatac tattgaagtt
4021 tctctggtac tcaaaggtta tgtagttaat aaattttcat taacaaaaaa

```

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 977 amino acids
- (B) TYPE: protein
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```

1 MSAGPGCEPC TKRPRWDAAA TSPPAASDAR SFPGRQRRVL DSKDAPVQFR VPPSSSGCAL
61 GRAGQHRGSA TSLVFKQKTI TSWMDTKGIK TVESESLHSK ENNNTREESM MSSVQKDNFY
121 QHNMEKLENV SQLGFDKSPV EKGTYLQKH QTAAMCKWQN EGPHERLLE SEPPAVTLVP
181 EQFSNANVDQ SSPKDDHSDT NSEESRDNDQ FLTHVKLANA KQTMEDQGR EARSHQKCGK
241 ACHPAEACAG CQOEETDVVS ESPLSDTGSE DVGTLGNAN RLNRQESSLG NSPPFEKESE
301 PESPMVDNS KNSCQDSEAD EETSPGFDEQ EDSSSAQTAN KPSRFQPREA DTELKRSSA
361 KGGEIRLHFQ FEGGESRAGM NDVNAKRPGS TSSLNVECRN SKQHGRKDSK ITDHFMVRVPK
421 AEDKRKEQCE MKHQTERKI PKYIPPHLSP DKKWLGTPIE EMRRMPRCGI RLPPLRPSAN
481 HTVTIRVDLL RIGEVKPFPP THFKDLWDNK HVKMPCEQN LYPVEDENGE RAAGSRWELI
541 QTALLNRLTR PQNLKDAILK YNVAYSKKWD FTALIDFWDK VLEEAEQHL YQSILPDMVK
601 IALCLPNICT QPIPLLKQKM NHSITMSQEQ IASLLANAF CTFPRRNAKM KSEYSSYPDI
661 NFNRLFEGRS SRKPEKLKTL FCYFRVTEK KPTGLVTFTR QSLEDFPEWE RCEKLLTRLH
721 VTYEGTIEGN GQGMLOVDFA NRVGGGVTS AGLVQEEIRF LINPELIVSR LFTEVLHDNE
781 CLITGTEQY SEYTYAETY RWARSHEDRS ERDDWQRRTT EIVAILDALHF RRYLDQFVPE
841 KIRRELNKAY CGFLRPGVSS ENLSAVATGN WCGAFGGDA RLKALIQILA AAVAERDVVY
901 FTFGDSELMR DIYSMHTFLT ERKLTVEGVY KLLRLRYNEE CRNCSTPGPD IKLYPFIYHA
961 VESCTQTTNQ PGQRTGA

```

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4069 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```

1 ggcgtctggg aagtgaggag cgtctctgcc tggcagaggc tgcaatctct gcactttggg
61 gggccaaggc aggcgctgag aaggacgcgc agtccatctc tctcaggtta gtgaaatgag
121 gctctccgcg gggccggccc ggggacagtg cgctgctggt cccagcatga atgccccccc
181 cggctgtgaa ccctgcacca aagcgacccg ctggggcgcc gctacaactt cgcggctgct
241 ttcggacgcc cggagcttct cgagcaggca gaggcgctc ctgcacccca aggacgctca
301 cgtgcagttc aggggtccac cgtctctgcc agcctgcgctc ccagggcagg cgggacagca

```

```

361 cagaggcagc gccacctcgc ttgttttcaa acaaaagact attaccagtt ggatggacac
421 taaaggaatc aagacagcgg aatcagaaa tttggatagt aaagaaaaca acaatacaag
481 aatagaatcc atgatgagtt ctgtacaaaa agataacttt taccaacata atgtagaaaa
541 attagtaaat gtttctcagc taagtcttga taagtcactc actgaaaaaa gtacacagta
601 tttgaaccag catcagactg cagcaatgtg taagtggcaa aatgaaggga aacacacgga
661 gcagcttttg gaaagtgaac ctcaaacagt aaccctggta ccagagcagt ttagtaatgc
721 taacattgat cggtcacctc aaaatgatga tcacagtgac acagatagtg aagagaatag
781 agacaatcaa cagttttctc caactgtaaa gcttgcaaat gcaaagcaga ctacggaaga
841 tgaacacgcc agagaagcca aaagccacca gaagtgcagc aagtcttgcc atcctgggga
901 agactgtgca agttgtcagc aagatgagat agacgtgggtg ccaaagagtc cattgtcaga
961 tgttggctct gaggatgttg gtactgggtc aaaaaatgac aacaaattga ttagacaaga
1021 aagttgccta ggaaattctc ctccatttga gaaggaaagt gaacccgaat caccgatgga
1081 tgtggataat tctaaaaata gttgtcaaga ctcagaagca gatgaggaga caagtccagg
1141 ttttgatgaa caagaagatg gtagttcttc ccaaacagca aataaacctt caaggttcca
1201 agcaagagac gctgacattg aatttaggaa acggtactct actaagggcg gtgaagttag
1261 attacatttc caatttgaag gaggagagag tcgcactgga atgaatgatt taaatgctaa
1321 actacctgga aatatttcta gcctgaatgt agaatgcaga aattctaagc aacatgaaa
1381 aaaggattct aaaatcacag atcatttgat gagactgccc aaagcagagg acagaagaaa
1441 agaacagtgg gaaaccaaac atcaaagaac agaaaggaag atccctaaat acgttccacc
1501 tcacctttct ccagataaga agtggttggt aactccattt gaggagatga gaagaatgcc
1561 tgggtgtggg atccggctgc ctctcttgag accatctgcc aatcacacag taactattcg
1621 cgtagatctt ttgcgagcag gagaagttcc taaacctttt ccaacacatt ataaagattt
1681 gtgggataac aagcatgtta aaatgccttg ttcagaacaa aatttgtacc cagtggaga
1741 tgagaatggt gagcgaactg cggggagccg gtgggagctc attcagactg cacttctcaa
1801 caaatattaca cgaccccaaa acttgaagga tgctattctg aaatacaatg tggcatattc
1861 taagaaatgg gactttacag ctttgatcga tttctgggat aaggtacttg aagaagcaga
1921 agctcaacat ttatatcgat ccatcttgcc tgatatggtg aaaattgcac tctgtctgcc
1981 aaatatttgc acccagccaa taccactcct gaaacagaag atgaatcatt ccatcacaat
2041 gtgcgaggaa cagattgcc a gtcttttagc taatgctttc ttctgcacat ttccacgacg
2101 aaatgctaag atgaaatcgg agtattctag ttaccagac attaaactca atcgattgtt
2161 tgagggaagt tcatcaagga aaccggagaa acttaaaacg ctcttctgct actttagaag
2221 agtcacagag aaaaaaccta ctgggttggt gacatttaca agacagagtc ttgaagattt
2281 tccagaatgg gaaagatgtg aaaaaccctt gacacgattg catgtcactt acgaaggtac
2341 catagaagaa aatggccaag gcatgtctaca ggtggatttt gcaaactcgt ttgttggagg
2401 tgggtgtaac agtgcaggac ttgtgcaaga agaaatccgc tttttaatca atcctgagtt
2461 gattatttca cggctcttca ctgaggtgct ggatcacaat gaatgtctaa ttatcacagg
2521 tactgagcag tacagtgaat acacaggcta tgctgagaca tatcgttggt cccggagcca
2581 cgaagatggg agtgaaaggg acgactgcga gcggcgctgc actgagatcg ttgccatcga
2641 tgctcttcac ttccagacgt acctcgatca gtttgtgctt gagaaaatga gacgcgagct
2701 gaacaaggct tactgtggat ttctccgtcc tggaagttct tcagagaatc tttctgcagt
2761 ggccacagga aactggggct gtgggtgcctt gccaggttaa aagccttaat
2821 acagatattg gcagctgctg cagctgagcg agatgtggtt tatttcacct ttggggactc
2881 agaattgatg agagacattt acagcatgca cattttcctt actgaaagga aactcactgt
2941 tggagatgtg tataagctgt tgctacgata ctacaatgaa gaatgcagaa actgttccac
3001 ccttggaaca gacatcaagc tttatccatt catataccat gctgtcagat cctgtgcaga
3061 gaccgctgac cattcagggc aaaggacagg gacctgagga gccgagcgaa tagcatctcc
3121 tcccacctcc caccagagac gtccctgtttg agctgtcagg tgtaatatat gaattgactt
3181 aagttaatat aaatgtgtac ataatccaca tttgtagtca aggacgcaat ctcttcaca
3241 catgtgcagt tgtcagttgg tacatctaaa ctccctccat cctgactcac gtggacttag
3301 atatgttttg tttctatttt cttctatttt agtttttcat tctttgatgt ttatttcttt
3361 tgtccatcag atctcttgtg aaatcccatg gaaggttggt ctcagctgtc ggggtctctt
3421 cttcctgccc atattattata ccagttgctt ctgcagcccg cagatgocca gogatgccc
3481 ggaaacaagt tgaaatccca ggaatctctt taactgattt tgctaaaaat ctccctgtga
3541 gccttccact caactcttaa tatgcttgca ttgtttaagt ttttaaattc tgaaaattaa
3601 taattagggt ttttttcata tgtgttgcac aatgcaaacc tcttaggtta aaatagtttc
3661 tttattttaag atagaataat ttccagaaat tgtacttttg aggtatcatt tttatctgta
3721 atggttttgt tgtctttttt cctctgatca gtattttttt ataccagttt tggagactgc
3781 ctgagatgaa aggaaatgtg gaataaaaag aggttttctt aaagaaaaca gaagtaataa
3841 gattccaaga gaattgaaga ttttttttgt ttcttggta ctttttctt tttaaattag
3901 gactaatgtt tcttttgtgg tgcttgaggc atattcatat aaccaagtt tgagaactgg
3961 gaacttcoatg ctgatttgta catattgaag tttctctggt attcaaaggt tatatagtga
4021 atgaattttc attaataaat cactttgtca gaaaaaaaaa aaaaaaaaaa

```

(2) INFORMATION FOR SEO ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 976 amino acids  
(B) TYPE: protein  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

|     |            |            |            |            |             |            |
|-----|------------|------------|------------|------------|-------------|------------|
| 1   | MNAGPGCEPC | TKATRWGAAT | TSPAASDARS | FPSRQRRVLD | PKDAHVQFRV  | PPSSPACVPG |
| 61  | QAGQHRGSAT | SLVFKQKIT  | SWMDTKGIKT | AESESLDSKE | NNNTRIESMM  | SSVQKDNFYQ |
| 121 | HNVEKLVNVS | QLSLDKSLTE | KSTQYLNQHQ | TAAMCKWQNE | GKHTEQLLLES | EPQTVTLVPE |
| 181 | QFSNANIDRS | PQNDHDSDTD | SEENRDNQOF | LTTVKLANAK | QTTDEHARE   | AKSHQKCSKS |
| 241 | CHPGDECASC | QQEIDVVPK  | SPLSDVGSED | VGTSGKNDNK | LIRQESCLGN  | SPPEKSESEP |
| 301 | ESPMVDVNSK | NSCQDSEADE | ETSPGFDEQE | DGSSSQTANK | PSRFQARDAD  | IEFRKRYSTK |
| 361 | GGEVRLHFQF | EGGESRTGMN | DLNAKLPGNI | SSLNVECRNS | KQHGGKDSKI  | TDHMLRLPKA |
| 421 | EDRRKEQWET | KHQRTERKIP | KYVPPHLSPD | KKWLGTPIEE | MRRMPCGIR   | LPLLRPSANH |
| 481 | TVTIRVDLLR | AGEVPKPFPT | HYKDLWDNKH | VKMPCEQNL  | YPVEDENGER  | TAGSRWELIQ |
| 541 | TALLNKFTRP | QNLKADAILY | NVAYSKKWDF | TALIDFWDKV | LEEAEAOHLY  | QSILPDMVIK |
| 601 | ALCLPNICTQ | PIPLLKQKMN | HSITMSQEIQ | ASLLANAFCC | TFPRRNAKMK  | SEYSSYPDIN |
| 661 | FNRLFEGRSS | RKPEKLTFLF | CYFRRVTEKK | PTGLVTFTRQ | SLEDFPEWER  | CEKPLTRLHV |
| 721 | TYEGTIEENG | QGMLQVDFAN | RFVGGGVTS  | GLVQEEIRFL | INPELIISRL  | FTEVLHDNEC |
| 781 | LIITGTEQYS | EYTGXAETYR | WSRSHEDGSE | RDDCERRCTE | IVAI DALHFR | RYLDQFVPEK |
| 841 | MRRELNKAYC | GFLRPGVSSE | NLSAVATGNW | GCGAFGGDAR | LKALIQILAA  | AAAERDVVYF |
| 901 | TFGDSELMRD | IYSMHIFLTE | RKLTVGDVYK | LLRYYNEEC  | RNCSTPGPDI  | KLYPFIYHAV |
| 961 | ESCAETADHS | GORTGT     |            |            |             |            |

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3814 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

|      |            |            |             |             |            |            |
|------|------------|------------|-------------|-------------|------------|------------|
| 1    | gggggactgt | gtgctgcggg | tcccagcatg  | agtgcggggc  | ccggctggga | gccctgcacg |
| 61   | aaagcgcgct | ggggcgccgc | tggaacttct  | gcgcgcgactg | cctcggactc | ccggagcttc |
| 121  | cctggcaggc | agaggcggtg | tctcgacccc  | aaggacgctc  | ccgtccagtt | cagggtccct |
| 181  | ccgtcctcgc | cagcctcgct | ctcggggcgg  | gcgggaccgc  | acagaggcaa | cgccacctcg |
| 241  | tttgttttca | aacaaaagac | tattactact  | tgaattggata | ctaaaggacc | caagacagct |
| 301  | gaatcagaaa | gtaaagaaaa | caacaataca  | tggaattgact | ccatgatgag | ttctgtgcag |
| 361  | aaagataact | tttaccacaa | taaggtggaa  | aaattggaaa  | atgttcctca | gctaaatctt |
| 421  | gataaatcac | ccacagaaaa | gagttcacag  | tatttgaacc  | aacagcagac | tgcgagtgtg |
| 481  | tgcaagtggc | agaatgaagg | gaagcatgca  | gaacagcttt  | tggcaagtga | gcctcccgcg |
| 541  | gggactccgc | taccaaagca | gcttagtaat  | gctaactattg | gtcagtcacc | ccacactgat |
| 601  | gaccacagtg | acacagatca | tgaagaagac  | agagacaact  | agcagtttct | tacacctata |
| 661  | aaacttgcaa | atacaaaagc | aacagtagga  | gatgggcagg  | ccagaagcaa | ctgtaagtgc |
| 721  | agtggatctc | gccagtctgt | gaaagactgt  | acaggctgtc  | aacaggagga | ggtggatgtg |
| 781  | ctaccagaga | gtcctttgtc | agatgttggt  | gccaggaca   | ttggaactgg | acaaaaaat  |
| 841  | gacaacaaat | tgactggaca | agaaagcagc  | ctaggtgatt  | cgctccatt  | tgagaaagaa |
| 901  | agtgagctcg | agcaccaat  | ggatgtgagc  | aactcgagaa  | acagttgtca | agattcagaa |
| 961  | gcagatgaag | aaacaagttc | agcttttgat  | gagcaagatg  | atcgttcttc | ccaaacagca |
| 1021 | aataaacttt | caagttgcca | agcaagagaa  | gctgatggcg  | atcttaggaa | acggtatttg |
| 1081 | actaagggaa | gtgaagttag | attgcatttc  | caatttgaag  | gagaaaataa | tgctgggacc |
| 1141 | agtgacttaa | atgccaagcc | atctggaaac  | tcttctagcc  | ttaatgtaga | gtgtagaagt |
| 1201 | tccaagcagc | atggaaaaag | ggattctaaa  | attacagatc  | atttcatgag | aatttcaaag |
| 1261 | tcagaggaca | gaagaaaaag | acaattgtgaa | gtcacagatc  | aaagaacaga | aaggaagatt |
| 1321 | ccaaaataca | tcccacttaa | ctccctcca   | gagagaagtt  | ggctgggaac | tctattga   |

```

1381 gaaatgagaa aaatgcctcg gtgtgggato catttgcott cottaagacc atctgcaagt
1441 cacacagtga ctgttcgggt agaccttctg agagcaggag aggttcgaa accttttcca
1501 acacattaca aagatttgtg ggataacaaa catgtgaaaa tgccttggtc ggaacaaaaa
1561 ttgtaccctg tggaagatga gaatggtgag cgaactgcag ggagtaggtg ggagctcatt
1621 cagactgcac ttctcaacaa attcacacga cccagaaact tgaaggatgc gattctgaaa
1681 tacaatgtgg catattctaa gaaatgggac tttacagctt tgggttgattt ctgggataag
1741 gtacttgaag aagcagaggc ccaacattta tatcagtcca ttttacctga catggtgaaa
1801 attgcactct gtctgccaaa tatttgcacc cagccaatac cactcctgaa acagaagatg
1861 aatcattctg tcacgatgtc acaggaacag atcgccagtc ttttagctaa tgctttcttc
1921 tgcacatttc cccgacggaa tgccaagatg aaatcggagt attctagtta ccagacattt
1981 aacttcaatc ggttgtttga aggacgttca tcaaggaaac cagaaaaact gaaaacactc
2041 ttctgtctact ttccaagagt cacagagaaa aaacctacag gattggtgac atttacaaga
2101 cagagtcttg aagattttcc agaatgggaa aggtgtgaaa agcctctgac acgcttacac
2161 gtcacttacg aggggtacat agaaggcaac ggccgaggca tgctacaggt ggattttgca
2221 aatcggtttt ttggaggtgg tgtgactggt gcgggacttg tacaagaaga aatcagattt
2281 ttaatacaatc ctgaattgat tgtttcacgg ctgttcactg aggtgctgga tcacaatgag
2341 tgtcttatta tcacaggtac tgaacagtac agtgaataca caggctatgc tgaacttat
2401 cgttggggccc gaagccatga agatgggagt gaaaaggacg attggcagcg gcgctgcacg
2461 gagatcggtg ccattgacgc acttcacttc agacgctacc togatcagtt tgtgcctgag
2521 aaagtgaagc gtgagcttaa caaggcttac tgcggattcc tccgtcctgg agttccttct
2581 gaaaatcttt ctgcagtggc caggggaaac tggggctgtg gtgccttttg gggtagcgt
2641 agattaaaaa ccttaataca gatcctggca gctgctgcgg ctgaacgtga cgtggtttat
2701 ttcacctttg gggactcaga gttgatgaga gacatttaca gcatgcacac ttctcttacc
2761 gagaggaagc tggatgttgg aaaagtgtac aagttattgc ttagatacta caatgaagaa
2821 tgcagaaact gttccacccc tggaccagac atcaagcttt atccattcat ataccatgct
2881 gttgagtcaa gtgcagagac cactgacatg ccaggacaga aggcaggcac ctgaggaaca
2941 agtgactagg acctcctctc aaagagacat cctatttgaa atgtggggtg tgatgtctga
3001 attgactgaa tctgatctaa gtgtgtatat aatccacatt tgtaatcaag gatgcagctc
3061 cttctgcata tgcagttggt tcttgttcat cctggtggac atgcctttag acatggcttc
3121 ttcaattttt cttctccttc agtctttatt ctttgatttt ttttttccaa cttgatttct
3181 tgggaaaact caagaaagg tgcactcagc ttctagatct ttctcttctt gtctgtgtgt
3241 tgtccagact gctttggtgg ctagcagata ccacacactc tggaggaagt tacaaatcca
3301 gaaatctgag tttgctgcag atttacctgt gagcttctca ctcccaaccc ttgttaggct
3361 tgtgttgtct acattttcaa ttttggaagt tgaagttttt cttatgttac ttaatgctag
3421 tatcttttag gctaaaaacta ttttctattt aaggcagact aatttccagt ttctcttttg
3481 aaacatcatc cctataagta acggtttttt tcgtcctttt tccccagcg ctatttttaga
3541 agctggccaa gaggaagaa aatgtagaat aaaaggattt tctcgggatg ctataaagaa
3601 gccaggttca agagcgttgg ggtttttgtt tttttcaaga cttgtttttc ctttgcagct
3661 aggggtgagt cttgttctgt ggtgctgagg gcatagtcct gtaaccaaag gtctttgctg
3721 gagacttgat gctgatttgt acatattgaa gtttctctgg caggaaatat tagagttaat
3781 aaatttcatt aataaatcat ttgtcagaaa aaaa

```

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 968 amino acids

(B) TYPE: protein

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

```

1   MSAGPGWEPC TKARWGAAGT SAPTASDSRS FPGRQRRVLD PKDAPVQFRV PPSSPACVSG
61  RAGPHRG NAT SFVFKQKTIT TWMDTKGPKT AESESKENNN TRIDSMSSSV QKDNFYPHKV
121 EKLENV PQLN LDKSPTEKSS QYLNQQQTAS VCKWQNEGKH AEQLLASEPP AGTPLPKQLS
181 NANIGQ SPHT DDHSDTDHEE DRDNQQFLTP IKLAN TKPTV GDGQARSNCK CSGSRQSVKD
241 CTGCQQ EEVD VLPESPLSDV GAEDIGTGPK NDNKLTGQES SLGDSPPFEK ESEPESPMDV
301 DNSRNS CQDS EADEETSPVF DEQDDRSSQT ANKLSSCQAR EADGDLRKRY LTKGSEVRLH
361 FQFEGEN NAG TSDLNAKPSG NSSSLNVECR SSKQH GKRD S KITDHFMRI S KSEDRRKEQC
421 EVRHQR TERK IPKYIPP NLP PEKKWLGTPI EEMRKMPRCG IHLPSLRPSA SHTVTVRVDL
481 LRAGEVP KPF PTHYKDLWDN KHKVMPCSEQ NLYPVEDENG ERTAGSRWEL IQTALLNKFT

```

|     |            |            |            |            |            |            |
|-----|------------|------------|------------|------------|------------|------------|
| 541 | RPQNLKDAIL | KYNVAYSKKW | DFTALVDFWD | KVLEEAEAOH | LYQSILPDMV | KIALCLPNIC |
| 601 | TQPIPLLKQK | MNHSVTMSQE | QIASLLANAF | FCTFPRRNAK | MKSEYSSYPD | INFNRLFEGR |
| 661 | SSRKPEKLKT | LFCYFRRVTE | KKPTGLVTFE | RQSLEDFPEW | ERCEKPLTRL | HVTYEGTIEG |
| 721 | NGRGLQVDF  | ANRFVGGGVT | GAGLVQEEIR | FLINPELIVS | RLFTEVLDHN | ECLIITGTEQ |
| 781 | YSEYTGAAET | YRWARSHEDG | SEKDDWQRRR | TEIVAIDALH | FRRYLDQFVP | EKVRRELNKA |
| 841 | YCGFLRPGVP | SENLSAVATG | NWGCFAFGGD | ARLKALIQIL | AAAAAERDVV | YFTFGDSELM |
| 901 | RDIYSMHTFL | TERKLDVGKV | YKLLRLRYNE | ECRNCSTPGP | DIKLYPFIYH | AVESSAETTD |
| 961 | MPGQKAGT   |            |            |            |            |            |

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2781 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

|      |             |             |             |             |             |             |
|------|-------------|-------------|-------------|-------------|-------------|-------------|
| 1    | tcgaagtgtg  | tggatatttat | aaagtgcgat  | attcatcaca  | gctatcgctc  | atccccaaaa  |
| 61   | caccggtatg  | caagaattca  | ggtcacacatt | gattttttcog | atattccaaa  | aggtttacca  |
| 121  | atctacggca  | aatcgccgca  | gagcaagtgc  | atccgtgctg  | accaatcgac  | tcggcaaggc  |
| 181  | tttgtgtctta | aactgcgcca  | ggatgtcgaa  | gtcgccggat  | ggcgggattt  | ccgaaataga  |
| 241  | aacggaggag  | gagccggaaa  | atctggcgaa  | ctccctagat  | gattcgtggc  | gtggagtttc  |
| 301  | catggaggct  | atacatcgta  | atcggcagcc  | tttcgaattg  | gagaatttgc  | caccagtgc   |
| 361  | tgcgggcaat  | ctccaccggg  | ttatgtacca  | gctgccaatt  | cgtgaaacac  | cgccacgccc  |
| 421  | ctacaaatca  | ccgggaaaagt | gggactccga  | gcatgtgcgt  | ctgcctgtg   | cgcccagtc   |
| 481  | gaaatatccg  | agggagaatc  | cggatggcag  | caccaccatc  | gattttcgct  | gggaaatgat  |
| 541  | cgaacgagcc  | cttctgcagc  | ccataaagac  | gtgtgaggaa  | ctgcaggcgg  | cgataatatc  |
| 601  | atataatacc  | acctataggg  | atcagtggca  | ctttcgtgcc  | cttcataaac  | ttctcgacga  |
| 661  | ggaactggac  | gagagcgaaa  | cacgggtttt  | cttcgaggat  | ctattgcgcg  | gcattatccg  |
| 721  | attggcattg  | cggctaccgg  | aattgattca  | atcgccagtt  | ccgctgctca  | agcaccacaa  |
| 781  | gaacgcctca  | ttgagcctga  | gccaaacagca | gatctcctgc  | ctggttgcca  | atgccttctt  |
| 841  | gtgcacgttt  | ccccgaagaa  | acaccctcaa  | gaggaaagtc  | gagtaacagca | cttttccaga  |
| 901  | catcaacttt  | aacaggcttt  | accaatcgac  | gggaccggca  | gttctggaga  | agcttaaatg  |
| 961  | cattatgcac  | tattttcgtc  | gcgtgtgtcc  | cacagagcgg  | gatgccagca  | atgtgcccac  |
| 1021 | cgggtgtggt  | acctttgttc  | gtcggagcgg  | attgcoggaa  | catctgatcg  | actggagcca  |
| 1081 | aagtgcggcg  | ccgttggttg  | atgtgccatt  | gcacgtggat  | gccgagggaa  | caatcgagga  |
| 1141 | tgagggcatt  | ggactgctgc  | aagttagactt | tgccaacaaa  | tatttgggtg  | gcggtgtctt  |
| 1201 | gggacatggc  | tgcgttcagg  | aggagatacg  | ctttgttatc  | tgtccggagc  | tatttggtgg  |
| 1261 | taaactcttt  | acggagtgtc  | tgcgaccatt  | cgaggccctg  | gtgatgttgg  | gcgccgaaag  |
| 1321 | gtatagtaac  | tatacgggat  | atgccggaag  | cttcgagtgg  | tccggcaact  | ttgaggattc  |
| 1381 | aacgccaaaga | gatagctcag  | gtcgtcgaca  | aacggccatt  | gtggcaatcg  | atgccctaca  |
| 1441 | ttttgcccag  | tcacatcatc  | aatatcgca   | ggatctcatg  | gaaagggagc  | tgaacaaggc  |
| 1501 | gtacattgga  | tttgttcaact | ggatggtgac  | gccgccaccg  | ggtgtggcaa  | ctggttaactg |
| 1561 | gggttgccgc  | gcattcggcg  | gtgactccta  | tctgaaagcc  | ctgctgcaac  | ttatggtctg  |
| 1621 | cgcccagttg  | ggcagacctt  | tggcctaact  | tacctttgga  | aatgtggagt  | ttagggatga  |
| 1681 | ttttcatgaa  | atgtggctgt  | tgtttcgaaa  | tgacgggact  | acggtgcagc  | agctttggag  |
| 1741 | tatttttaagg | togtacagta  | ggcttattaa  | ggagaagagc  | tccaaggagc  | cgctgtagaa  |
| 1801 | taaggcatoc  | aaaaagaagc  | tatatgattt  | tattaaagag  | gaacttaaga  | aggtcagaga  |
| 1861 | tgtgcccggg  | gagggagcat  | ccgccgaagc  | tggaagctct  | agagtagctg  | gattaggcga  |
| 1921 | aggaaaatca  | gaaacatcag  | cgaaatcctc  | gccagaactc  | aacaagcaac  | ccgcccgacc  |
| 1981 | gcaaatcacc  | ataacgcaac  | aaagtaccga  | tctattgccc  | gcgcaattat  | cgcaagataa  |
| 2041 | ctctaattct  | tcggaagatc  | aggcccttct  | tatgctgtcg  | gacgatgagg  | aggccaatgc  |
| 2101 | catgatggag  | gocgctagtc  | tggaggctaa  | aagcagcgta  | gaaataagca  | acagcagcac  |
| 2161 | aacgtccaaa  | acgagcagta  | cagccacgaa  | atcaatgggt  | tcaggtggcc  | gccagttgag  |
| 2221 | tctgtctogag | atgctggaca  | cccattatga  | aaagggttcg  | gctctgaaga  | ggccacgaaa  |
| 2281 | atcaccacaac | tgcagcaagg  | ctgagggttc  | agcaaagagt  | cgtaaggaga  | tcgatgtgac  |
| 2341 | cgacaaggac  | gaaaaggacg  | atattgttga  | ctaggtgata  | ttgcactaca  | ggattgttac  |
| 2401 | tgcccccaaa  | aattgaagag  | gtataaaatg  | tattgtagat  | aactttaagg  | acatatttag  |
| 2461 | ggcatttttaa | agtaggatca  | ttgtaagtgc  | aataaagtga  | aatttttttt  | tttttttaat  |

2521 tatactatttc taatctgcaa agacaatddd actgttaaat ttgtataaca ttogaattaa  
 2581 ttaataataat ttgttatatc atgcaaactc agctttttatt atgcgaaatt tgtagttaaaa  
 2641 gccagtaaaag tttctttttta ttttaaccgaa accttttggt tatttttattt gaccacaaca  
 2701 agaacatcaa caacaacaac caccgaaaaaa aagcgaatat atatttggtt gttcgtatat  
 2761 atatatatat ctaagcagat c

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 768 amino acids
- (B) TYPE: protein
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

1 MQEFRSHLIF PIFQKVYQST ANRRRASASV LTNRLGKALC LNCARMSKSP DGGISEIETE  
 61 EEPENLANSL DDSWRGVSM EAIHRNRQPF LENLPPVTAG NLHRVMYQLP IRETPPRPYK  
 121 SPGKWDSEHV RLPAPESKY PRENPDGSTT IDFRWEMIER ALLQPIKTCE ELQAAIISYN  
 181 TTYRDQWHFR ALHQLLDEEL DESETRVFFE DLLPRIIRLA LRLPDLIQSP VPLLKHHKNA  
 241 SLSLSQQQIS CLLANAF LCT FPRRNTLKRK SEYSTFPDIN FNRLYQSTGP AVLEKLK CIM  
 301 HYFRRVCPT E RDASNVP TGV VTFVRRSGLP EHLIDWSQSA APLGDVPLHV DAEGTIEDEG  
 361 IGLLQVDFAN KYLGGGV LGH GCVQEEIRFV ICPELLVGKL FTECLRPFEA LVMLGAERY S  
 421 NYTGYAGSFE WSGNFEDSTP RDSSGRRQTA IVAIDALHFA QSHHQYREDL MERELNKAY I  
 481 GFVHVWMTTP PGVATGNWGC GAFGGDSYLK ALLQLMVCAQ LGRPLAYYTF GNVEFRDDFH  
 541 EMWLLFRNDG TTVQQLWSIL RSYSLRIKEK SSKEPRENKA SKKKLYDFIK EELKKVRDVP  
 601 GEGASAEAGS SRVAGLGEGK SETSAKSSPE LNKQPARPQI TITQQSTDLL PAQLSQDNSN  
 661 SSEDQALLML SDDEEANAMM EAASLEAKSS VEISNSSTTS KTSSTATKSM GSGGRQLSLL  
 721 EMLDTHYEKG SASKRPRKSP NCSKAEGSAK SRKEIDVTDK DEKDDIVD

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2181 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

1 ATGAGCAAGA AGTTTATCGA ACTGGGTGAT CCTGTCACTC AAGACGAGAA  
 51 AGACTACGAA GACTATGTCG GAGTTGGTTT CGCGCATCAA GTCCCGACAA  
 101 TGAAAAGGCG GAAGTTGACA GAACATGGAA ATACTACAGA ATCAAAAAGAA  
 151 GATCCTGAAG AGCCAAAAAG CCGTGACGTA TTTGTCTCCT CGCAGTCAAG  
 201 TGATGAGAGT CAAGAAGATT CGGCTGAAAA TCCGGAGATC GCTAAAGAAG  
 251 TGTCAGAAAA TTGTGAAAAT CTGACAGAAA CTCTCAAAAT TTCTAATATT  
 301 GAGAGTTTGG ACAATGTTAC TGAAAGATCT GAACACACTC TTGATAATCA  
 351 CAAAAGTACT GAACCAATGG AAGAAGATGT AAACAACAAG TCCAATATTG  
 401 ACGTTGCGAT TAATTCTGAC GAGGATGATG AACTTGTTCT GGAAGAGAAT  
 451 AATAAAGAAA TGAGGGATGG AGAACAAGTA CAACAGTTGT CACAGGATTT  
 501 ATTCGCTGAT GATCAAGAGC TAATTGAATA TCCAGGAATT ATGAAAGACA  
 551 CTACAACTCA ACTGGATATA ACAGATTCTG AAGTGGAGAC TGCTCAAAAA  
 601 ATGGAAATGA TTGAAGAAAC TGAAGCAGAT TCGACATTTG TAGGCGAGGA  
 651 TTCAAAAAGCT ACGAAAAGCT TGAGGACATC CAGTTCAAGT TTCCTGTCAA  
 701 CTGTTTCAAC ATGCGAAGCC CCTGCAAAAG GACGAGCAAG AATGTATCAA  
 751 AAAGAGTTGG AAAAGCATGT GATTGCATTT ACTGAGGGAA ATCTCACACT  
 801 ACAACCAGAT TTGAACAAAG TTGATCCCGA CAGAAACTAT CGATATTGTA  
 851 CAATTCCGAA CTTTCCAGCT TCCCAAGGAA AACTTCGAGA AGATAATCGA  
 901 TATGCCCCAA AAATCGTTTT GCCTCAAAGA TGGCGAGAAT TTGATTGAG  
 951 GGGCCGTTAG AGAGACTCAT ATTTCTATTT CAAACGTAAG CTCGATGGAT



```

1001 ATTTGAAATG CTACAAAACA ACTGGATATT TTATGTTTGT TGGACTTTTG
1051 CACAACATGT GGGAAATTTGA CCCAGACATC ACATATAAAC TGCCAGCACT
1101 GGAAATGTAT TACAAAGAGA TGTCGGAACT TGTGTTAGTA GAAGAGGTTT
1151 TGGAAAAATT TGCACGAGTT GCGCCGATCG CAAAAACTGC TGAAGATATT
1201 CTGCCAGAGC GAATTTATCG TCTTGTTGGT GACGTCGAAT CAGCTACCTT
1251 GAGCCACAAG CAATGTGCTG CACTTGTTGC GAGAAATGTTT TTTGCCCCGAC
1301 CGGACAGTCC TTTTCACTTC TGCCGAATTC TCTCGTCTGA TAAATCTATT
1351 TGTGTGGAGA AACTTAAATT CCTGTTCACT TATTTTCGACA AAATGTCAAT
1401 GGATCCACCG GATGGTGCCG TCAGTTTTAG ACTTACAAAA ATGGATAAAG
1451 ATACGTTCAA CGAAGAGTGG AAAGATAAAA AATTACGTTT TCTTCCTGAA
1501 GTTGAATTCT TTGATGAAAT GCTTATTGAA GACACAGCTC TCTGTACACA
1551 AGTTGATTTT GCGAACGAAC ATCTTGGTGG CGGAGTTTGA AATCATGGGT
1601 CTGTTTCAGGA GGAGATCCGT TTCTTGATGT GTCCAGAAAT GATGGTTGGA
1651 ATGTTGTTGT GCGAGAAAAA GAAACAACCTG GAAGCGATTG CAATTGTTGG
1701 AGCTTACGTT TTCAGTTCTT ATACTGGTTA TGGTCATACT CTAAAAATGGG
1751 CAGAACTTCA ACCAAATCAT TCTCGTCAGA ATACAAACGA ATTTTCGAGAT
1801 CGTTTTGGAC GTCTTCGGGT AGAAACTATT GCAATCGATG CAATTCTGTT
1851 CAAAGGATCA AAATTAGATT GTCAGACGGA GCAGTTAAAC AAAGCAAATA
1901 TCATTAGGGA AATGAAGAAA GCATCTATCG GATTCATGAG CCAGGGACCG
1951 AAATTCACAA ATATTCCAAT TGTTACTGGA TGGTGGGGAT GTGGAGCATT
2001 TAATGGGGAC AAGCCACTGA AGTTCATAAT CCAAGTAATT GCTGCCGGAG
2051 TCGCTGATCG TCCACTTCAT TTCTGTTTCA TGGGAGAACC CGAGCTTGCC
2101 GCAAAGTGCA AGAAAATTAT AGAACGAATG AAACAGAAGG ACGTAACACT
2151 TGGTAAGTCA TGTTTTTCAA TCTTCAGTTG A

```

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 726 amino acids
- (B) TYPE: protein
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

```

1 MSKKFIELDG PVTQDEKDYD DYVGVGFAHQ VPTMKRRKLT EHGNTTESKE DPEEPKSRDV
61 FVSSQSSDES QEDSAENPEI AKEVSENCEN LTETLKISNI ESLDNVTERS EHTLDNHKST
121 EPMEEDVNNK SNIDVAINSD EDELVLLEN NKEMRDGEQV QQLSQDLFAD DQELIEYPGI
181 MKDTTQLDI TDSEVETAQK MEMIEETead STFVGEDSKA TKTVRTSSSS FLSTVSTCEA
241 PAKGRARMYQ KELEKHVIAF TEGNLTLPD LNKVDPDRNY RYCTIPNFPA SQGKLREDNR
301 YGPKIVLPQR WREFDSRGRR RDSYFYFKRK LDGYLKYCT TGYFMFVGLL HNMWEFDPDI
361 TYKLPALEMY YKEMSELVGR EEVLEKFARV ARIAKTAEDI LPERIYRLVG DVESATLSHK
421 QCAALVARMF FARPDSPFSF CRILSSDKSI CVEKLKFLFT YFDKMSMDPP DGAVSFRLTK
481 MDKDTFNEEW KDKKLRLSLE VEFFDEMLIE DTALCTQVDF ANEHLGGGV LNHGSVQEEIR
541 FLMCPMMVG MLLCEKMKQL EAISIVGAYV FSSYTGYGHT LKWAELQPNH SRQNTNEFRD
601 RFGRLRVETI AIDAILFKGS KLDCQTEQLN KANIIREMKK ASIGFMSQGP KFTNIPIVTG
661 WWGCGAFNGD KPLKFIIQVI AAGVADRPLH FCSFGEPELA AKCKKIIERM KQKDVTLGKS
721 CFSIFS

```

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 amino acids
- (B) TYPE: polypeptide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein fragment

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

LFTEVLHDNE CLIITGTEQY SEYTGyaety R

- (2) INFORMATION FOR SEQ ID NO: 12:
- (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 29 amino acids
    - (B) TYPE: polypeptide
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein fragment
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

AYCGFLRPGV SSENLSAVAT GNXGCGAFG

- (2) INFORMATION FOR SEQ ID NO: 13:
- (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 11 amino acids
    - (B) TYPE: polypeptide
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein fragment
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

FLINPELIVS R

- (2) INFORMATION FOR SEQ ID NO: 14:
- (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 16 amino acids
    - (B) TYPE: polypeptide
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein fragment
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

IALXLPNIXT QPIPLL

- (2) INFORMATION FOR SEQ ID NO: 15:
- (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 17 amino acids
    - (B) TYPE: polypeptide
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein fragment
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

GAYCAYAAYG ARTGYTT

- (2) INFORMATION FOR SEQ ID NO: 16:
- (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 17 amino acids
    - (B) TYPE: polypeptide
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein fragment
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

CKRTANGTYT CNGCRTA

- (2) INFORMATION FOR SEQ ID NO: 17:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 24 bases  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: oligonucleotide  
  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

ATCATCACAG GTACTGAGCA GTAC

- (2) INFORMATION FOR SEQ ID NO: 18:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 24 bases  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: oligonucleotide  
  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

GCCTGTGTAT TCACTGTACT GCTC

- (2) INFORMATION FOR SEQ ID NO: 19:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 26 amino acids  
 (B) TYPE: polypeptide  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: polypeptide  
  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

EDKRKEQCEM KHQRTKIP KYIPPH

- (2) INFORMATION FOR SEQ ID NO: 20:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 26 amino acids  
 (B) TYPE: polypeptide  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: polypeptide  
  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

EDRRKEQWET KHQRTKIP KYVPPH

- (2) INFORMATION FOR SEQ ID NO: 21:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 26 amino acids  
 (B) TYPE: polypeptide  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: polypeptide  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

EDRRKEQCEV RHQRTKIP KYIPPN

- (2) INFORMATION FOR SEQ ID NO: 22:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 32 amino acids  
 (B) TYPE: polypeptide  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: polypeptide  
  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

HQVPTMKRRK LTEHGNTTES LLLKEDPPEP KS

- (2) INFORMATION FOR SEQ ID NO: 23:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 26 amino acids  
 (B) TYPE: polypeptide  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: polypeptide  
  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

EGKRKGDEVD GVDEVAKKKS KKEKDK

- (2) INFORMATION FOR SEQ ID NO: 24:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 26 amino acids  
 (B) TYPE: polypeptide  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: polypeptide  
  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

EGKRKGDEVD GTDEVAKKKS RKETDK

- (2) INFORMATION FOR SEQ ID NO: 25:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 26 amino acids  
 (B) TYPE: polypeptide  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: polypeptide  
  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

EGKRKGDEVD GIDEVTKKKS KKEKDK

- (2) INFORMATION FOR SEQ ID NO: 26:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 25 amino acids  
 (B) TYPE: polypeptide  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: polypeptide  
  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

EGKRKGEEVD GNVVAKKKS KEKEK

- (2) INFORMATION FOR SEQ ID NO: 27:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 26 amino acids  
 (B) TYPE: polypeptide  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: polypeptide  
  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

EGKRKADEV D GHSAATKKKI KKEKEK

- (2) INFORMATION FOR SEQ ID NO: 28:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 25 amino acids  
 (B) TYPE: polypeptide  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: polypeptide  
  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

EELPDTKRAK MELSDTNEEG EKKQR

- (2) INFORMATION FOR SEQ ID NO: 29:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 31 amino acids  
 (B) TYPE: polypeptide  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: polypeptide  
  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

EGVSSAKKAK IEKIDEEDAA SIKELTEKIK K

- (2) INFORMATION FOR SEQ ID NO: 30:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 28 bases  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: oligonucleotide  
  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

GCTGCGGGTC TCGACGATGA GTGCGGGC

- (2) INFORMATION FOR SEQ ID NO: 31:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 29 bases  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: oligonucleotide  
  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

GCGTCTAGAA TTCACCTGGC TCCTCAGGC

- (2) INFORMATION FOR SEQ ID NO: 32:  
 (i) SEQUENCE CHARACTERISTICS:  
   (A) LENGTH: 38 bases  
   (B) TYPE: nucleic acid  
   (C) STRANDEDNESS: single  
   (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: oligonucleotide  
  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

CCGGAATTCG GGTTTTTTGT TAATGAAAAT TTATTAAC

- (2) INFORMATION FOR SEQ ID NO: 33:  
 (i) SEQUENCE CHARACTERISTICS:  
   (A) LENGTH: 29 bases  
   (B) TYPE: nucleic acid  
   (C) STRANDEDNESS: single  
   (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: oligonucleotide  
  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

TCAGAGCAGA TGAATCGAG CAGTCCAGG

- (2) INFORMATION FOR SEQ ID NO: 34:  
 (i) SEQUENCE CHARACTERISTICS:  
   (A) LENGTH: 61 bases  
   (B) TYPE: nucleic acid  
   (C) STRANDEDNESS: single  
   (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: oligonucleotide  
  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

1 CCAATTTGAA GGAGGAATTC CCGCCGCCAC CATGAATGAT GTGAATGCCA AACGACCTGG  
 61 A

- (2) INFORMATION FOR SEQ ID NO: 35:  
 (i) SEQUENCE CHARACTERISTICS:  
   (A) LENGTH: 22 bases  
   (B) TYPE: nucleic acid  
   (C) STRANDEDNESS: single  
   (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: oligonucleotide  
  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

gaattcccgccg ccgccaccATG AA

- (2) INFORMATION FOR SEQ ID NO: 36:  
 (i) SEQUENCE CHARACTERISTICS:  
   (A) LENGTH: 674 bases  
   (B) TYPE: nucleic acid  
   (C) STRANDEDNESS: single  
   (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: cDNA  
  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

```

1   agaagaaaat ggccaaggca tgetacaggt ggatttttgca aatcgttttg ttggaggtgg
61  tgtaaccagt gcaggacttg tgcaagaaga aatccgcttt ttaatcaatc ctgagttgat
121 tatttcacgg ctcttcactg aggtgctgga tcacaatgaa tgtctaatta tcacaggtac
181 tgagcagtac agtgaatata caggctatgc tgagacatat cgttgggtccc ggagccacga
241 agatgggagt gaaagggacg actgcgagcg gcgctgcact gagatcggtg ccatcgatgc
301 tcttcacttc agacgctaacc tcgatcagtt tgtgcctgag aaaatgagac gcgagctgaa
361 caaggcttac tgtggtattc tccgtcctgg agtttcttca gagaatcttt ctgcagtggc
421 cacaggaaac tggggctgtg gtgccttttg ggggtgatgc aggttaaaag ccttaataca
481 gatattggca gctgctgcag ctgagcgaga tgtggtttat ttcacctttg gggactcaga
541 attgatgaga gacatttaca gcatgcacat tttccttact gaaaggaaac tcactgttgg
601 agatgtgtat aagctgttgc tacgatacta caatgaagaa tgcagaaact gttccacccc
661 tggaccagac atca

```

(2) INFORMATION FOR SEQ ID NO: 37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 200 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

```

1   aaaaatagtt gtcaagactc agaagcagat gaggagacaa gtccagggtt tgatgaacaa
61  gaagatggta gttcctccca aacagcaaat aaaccttcaa ggttccaagc aagagacgct
121 gacattgaat ttaggaaacg gtactctact aagggcgggtg aagttagatt acatttccaa
181 tttgaaggag gagagagtcg

```

(2) INFORMATION FOR SEQ ID NO: 38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29793 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

```

1   gatctcgaag taaaaactca cgcagaaaaga gtcctcctc ctttagcatg agaatccaac
61  tttgtaatga taacactggc aacatcaaca gtttgagaga aagcacgtgc ttgggcttca
121 caagcttgtc caatagaagc atccatcaca aaaacaacat tatctggtgt aactgcgttg
181 gaaacttgga gcatttcttc gaaaagtga gcttcttgct tgtgaogacc tgatgtatca
241 acaatgatga ttctgaacct ttcttgctgc aaaacaaata ttattaaacc atttttctgt
301 gataaattac cgtgaatttt tctactcctt cggcggcaat ttttacgggg tcaatttcag
361 agtatgatcc atagaaggga atacgagctt ttgtggcatt ttgctttaat tgatcaaaag
421 ctccagcacg gaatgtatcg gcacagatca gacatgtttt ccatcctttt ctttggtagt
481 aatacgccat ctgaacttga aaagtgttga aaagtgttga gaagtttact aattaaaaaa
541 tataatgttt gatggtgtgt gagctttcta ttgtaattca tggaaacgaac cttggtacaa
601 gtcgtagttt taccggaacc ttgaagacca acaaacatga aaacgttgcg acgtcctttt
661 gttggtgtga aaggagttac accaggatcc acaagcttca gcagttcatt gaatactgtc
721 ttctgaatgt accgacgttt gtttgctcct ccgaogatct cttogaaatt aatogctttt
781 ctgaaaatat ttattaaatt taaatcttaa atagcgtaaa aatttacttc acgttgtcct
841 taagttgctt tacaagacga atatgaacat cagattcaat aagagctgta cagacttctt
901 tcagcatcaa atccagctcc cctcatttga taacggtgct ctgacogagc tttcogatcg
961 catttcggat tttccgcccc aaatcggcc aaccattttt gaactgaaat ttgaaatgct
1021 ttaatttggt taagcataga attaaacgog ttttaaattg agagcaacct aaaaacagtt
1081 tggagaaaaa tcgataattc ttgtaggaga tttaagtoct gtgggtttct tcggcttctt
1141 aatcattttt tgacgacata gtggtatttc acaatagggt ttttcaagac acaacagatt
1201 tttcaciaag agtagagaag aaatggaaaa ctgtagattt cttctcgaag agccgagaaa
1261 ggcaagggtat tggaaagtta aaaaggtaat gtttctttat tcttttttca aaacaataat

```

```

1321 aaatggaaaa tatatatatta tagataacaa ttccagacag ttaaaatcac gtgaaaaatt
1381 caaatattcaa cacaaaaaatt gacgagtgga acccgggtgt tgcgccttga agagtaacgc
1441 ttgcgcggttt gacgattttta ttgacgcggt totggtgcat gcgggaaatt ttttattttc
1501 aactttttttc ctgttttggtt atcctttttt aattgaaattc tcatgatttg aaagctttga
1561 aaaatatttat tttgctcaaa aacatgcgtt ttgtaaaaca ttgattagat tcaaggcaat
1621 taatggatttt ttgcacgttc caaaaaaaag gaaattcatt ttttgaaaat tttgataatt
1681 taataatgaa aaatgttcca tagattttatt caatgccatc cttctctata atctcgaact
1741 tccgcacccct tcaactgtgg tagaggtatt tgcaataacca tatagtcgta ataataaact
1801 ttagtgaaca aatccaagac atcagctctt gagtaaatga atgattttata aaaactgctg
1861 attttctcgt aggaagaaaag agaatcagct aataatccgt cgttgtctat tctgtcaggc
1921 cgcttaaatg ttaaaaaata aaaacgtttt aagctaattt tgtatgtcta gaaactctaa
1981 ctcaacaagca tttctgcata cgccggatta gttggttttg caaaaagcga gtaatctaca
2041 aaagtgaatt tttgattcat ctctccattt tcacaaaacc aattttgtgg taogtatttc
2101 atatgatctt catccacttt ttttagttttt gaatgtattt gtgtgagttg tgtccagatt
2161 tgaataagat aacatctcag atccaaactg caattgaagc aagaacgatc tctctgaaa
2221 ttttatatga ccttaaaactt tatgtttcgtc tagtttcgtc gatatctgat cgttcaggtg
2281 tataggtatg tacatctcta ggtttatgtg ctacacgaaa atataatttg ttttacctaa
2341 cacacgcac cataaaatga tctacaaatc gttcaatttg atcctgtctt ggaaataata
2401 atttccaatt cgtaaagttt gcattcaact cattttctcg tttcaaactg togatattcg
2461 caaaatatgt tagtgaatca ctatcacaca ctctgaaaag cacaatatcc atattttcgt
2521 gttataatg aacctcacga ttcattcatta aatttctctt ggagcccgca taatacttgc
2581 tgcccaatta aagtatcagt ttcacagatt gcagttctat catttccgat agcctcaaat
2641 aagatttaatt cttaagcgag tgttctgac aatttaaata tttgatactc accgcaagtt
2701 tcttcgaaac ttgttcgaaa gctggaattt tagaatatcc tcaaaaactt ttttctcgc
2761 cctcatcaag ccataataag ttttgatcag caatatattc gaataaatta gtctctgata
2821 aatctcgtat cacaatcttt ttttctactc taaagaatac aattttgata agaataataa
2881 taattataat tataatagtt cgtcgtgag ttgatgaaga ccacataatt agtttaattg
2941 caagctatgc aacttggtga atactaatag gacttagcaa atcttatctt gaaacttttt
3001 cattcgaaag aaaaatgaga tcgaatctcg ttcaaactgt ggagtagtca gtttaagaaac
3061 ttgtttctag tttgtgagga gacactggag aacgtgaaag tattacccat acgcaatatt
3121 tttgcggcga aaaatacggg acccgggtctc gacacgacag tttttaaaac ttgtaaatag
3181 gtatgtaaaa gaaaacttta attttaaacg tgttgtttcg gaattttcat cgttttgtca
3241 tagttattct acaaataatt atttatgaaa aaaaaactaa aatataacta taataacacc
3301 tgaatattaa caaatcgatc gaaaaaaaac tatgaaaaaa atggatgaaa attccgcagc
3361 aacgagagtt tgaaatttca gtattcttta aaggcttacc gatttcaata aatagtgaac
3421 ctgaaaattg tagtttttaa actagttggt tagtatcatc aaatattcaa tcttcaaaa
3481 attcctcaat attaacgtat tttctctaatt tgtcttcatt atctaaaaaa aagttgcaat
3541 atatttttcc aggcagaaat agactttcac aaaacacatc gacacttoga atgagcaaga
3601 agttttatcga actgggtgat cctgtcactc aagacgttag ttatagtttt tattacttga
3661 acattatcat ctttttacag gagaaaagact acgaagacta tgtcggagtt ggtttcgcgc
3721 atcaagtccc gacaatgaaa aggcggaagt tgacagaaca tggaaatact acagaatcaa
3781 aagaagatcc tgaagagcca aaaagccgtg acgtatttgt ctcctcgcag tcaagtgatg
3841 agagtcaaga agattcggct gaaaatccgg agatcgctaa agaagtgtca gaaaattgtg
3901 aaaatctgac agaaaactctc aaaatttcta atattgagag tttggacaat gttactgaaa
3961 gatctgaaca cactcttgat aatcacaaaa gtactgaacc aatggaagaa gatgtaaaaa
4021 acaagtccaa tattgacgtt gcgattaatt ctgacgagga tgatgaactt gttctggaag
4081 agaataataa agaaatgagg gatggagaac aagtacaaca ggtcaggaaa ttttacaagt
4141 gaatgaaata agttaatcac caaaatgaat aaggacattt cccatcagaa aggtctctcg
4201 aatttttaggt gtaatgttaa ttttttgctg tagtttttcc cattgtttga aatttttgcc
4261 aaaattagtt attgcatacc cttcatgttt ttgaagattg tttaggaatg agaaaaacatt
4321 ttggacgctt ttattattag gacaccaaac actttttgtt gaaaaaacag ctcgttttaa
4381 aaaagctttt tcaaaaaaat ctgacgcaag gcttgtgaat tttcgttttc cctgatattt
4441 taaaatttct cctaaagttt tttgctaata tttttcgcta tgcgctaatt tactagtga
4501 tcaacaaaaa attttttttt tttcatagat tttttataag tttttgaaaa catagattta
4561 aaacttaaac ttaaattttg acaaggcgag aggaaaaaat taaaaattgc tgaacattca
4621 gatgccggtt acctattttt tgggttcaaaa atcccaatat tacgcgtctg ggttatagtc
4681 atttgccctt attaaattaa tgggtgttctt tggaaaagta agttctgttt tgttttcagc
4741 tttcacttc atcaaacgga aggaagggtt gattaaggaa agtaaacata ttttatgttg
4801 ttcttgtcac ttcctccatt tcgcaataat ataactcgag aaatatagaa ttttgttoga
4861 agttttcttt ttccttcaac attttaataa ttgttagtat taccagaaa aatagaaaaa
4921 atcgaagaaa tttgcaaaaa agcagacgta gaggctacgt acttcttaag cagccctt
4981 ttctttttaa tttgttcggt cgtaccgaga tccggtacct tattttacaa cgttttctgt

```



|      |             |             |             |             |             |             |
|------|-------------|-------------|-------------|-------------|-------------|-------------|
| 5041 | tccaaaaata  | ataatgtact  | gcagttgtca  | caggatttat  | tcgctgatga  | tcaagagcta  |
| 5101 | attgaatatc  | caggaattat  | gaaagacact  | acaactcaac  | tggatataac  | agattctgaa  |
| 5161 | gtggagactg  | ctcaaaaaat  | ggaaatgatt  | gaagaaactg  | aagcagattc  | gacatttgta  |
| 5221 | ggcgaggatt  | caaaagttag  | acaaaatcat  | tctgacaagg  | attcctgcga  | gcactcagtc  |
| 5281 | aagagcgagt  | cacggcaact  | cgggtccaaaa | ccattttctaa | ttagtaaact  | ctcaaaaaacc |
| 5341 | acaactaaat  | agcttaaaac  | ctttgtaaat  | tagcttattt  | ttgctaatta  | gcaatgattt  |
| 5401 | taagctaatt  | agttgtgggt  | tttgagaggt  | tactaattag  | aaatggtttt  | ggaccgagtt  |
| 5461 | gccgtgactc  | gctcttgact  | gagcacaagc  | aaacttttgt  | ggatgttgag  | aatcagcggc  |
| 5521 | aaagtggcac  | tactagttag  | gaagttgacg  | cagattctca  | gattaatttg  | gtaagacaaa  |
| 5581 | gaaaatataa  | atttttattac | ccagatgcat  | attttcatga  | ttctgatgca  | aaaaatacgg  |
| 5641 | taccgatct   | ggatactaca  | atttttgttaa | aatgcgaaaa  | ggttttgcacc | tttaaaaaaga |
| 5701 | actgcaattt  | caaacacttg  | ttgctgtgga  | ttgtttatcg  | gtttttaata  | ttttttggtg  |
| 5761 | agagtaaattg | agaaaagcga  | gttcccgcat  | tatctgtgtg  | cgattttggaa | tacagtactt  |
| 5821 | ttcaaagacg  | cacaccattt  | tgcatataac  | aaacatttgt  | cgtgtcgaaa  | ccgggtaccg  |
| 5881 | tgattttgca  | ttaaaagttg  | caaaatttca  | catagttttt  | ataattttag  | ctacgaaaa   |
| 5941 | tgtgaggaca  | ttcagttcaa  | gtttcctgtc  | aactgtttca  | acatgcgaag  | cccttgcaaa  |
| 6001 | aggacgagca  | agaatgtatc  | aaaaagagtt  | ggaaaagcat  | gtgattgcat  | ttactgaggg  |
| 6061 | aaatctcaca  | ctacaaccag  | atltgaacaa  | agttgatccc  | gacagaaact  | atcgatattg  |
| 6121 | tacaattcog  | aacttttocag | cttcccaagg  | aagtacgttg  | ttcaataaaa  | catactaggt  |
| 6181 | atataattaa  | ttattttcaga | acttcgagaa  | gataatcgat  | atggcccaaa  | aatcgttttg  |
| 6241 | cctcaaagat  | ggcgagaatt  | tgattcggta  | catttctatt  | gaatttaatta | tatactactt  |
| 6301 | actagaacaa  | ccatggagaa  | agaatgcgaa  | aaattgaatt  | ttaaaaacta  | atlttttaatt |
| 6361 | tttggctaaa  | ttttcagttt  | gaattttaatc | caaaatgaaa  | actgcgacca  | atcaatgact  |
| 6421 | tttcaaaatc  | actttttcaac | caatcaaacg  | gagtgtgctg  | ggctcgaaga  | cgctgattgg  |
| 6481 | ttcggaatg   | ggcgtgggtt  | ctcatttttg  | agggaaattca | aaaaaaggca  | tttggtcaca  |
| 6541 | gttgaaaatc  | atgtttttcaa | aagatgcatt  | ttttattcct  | ctctgatttt  | ttttgatttt  |
| 6601 | ctttttgtgt  | atltctgaat  | ttaaaggttg  | tgtagtgcga  | tttttttatt  | gctttattag  |
| 6661 | actcaaaatt  | ttctgaaaac  | gccaaatttc  | ataatgaaac  | ttcttgaaaa  | ctcttcagca  |
| 6721 | aaaagttatg  | acggctcaaa  | aaatggccta  | aaattagtta  | agattggaga  | tttgaccgac  |
| 6781 | ttgtcaatgt  | cgcagcggct  | ggaaacaatt  | ttttttgaaa  | tcaccgtcaa  | atlttaagta  |
| 6841 | tacaacttga  | ttatttttgcg | ttttaaactt  | tatttaggta  | tttaaaagtc  | gatggacggc  |
| 6901 | gagttttggc  | tcaaaaaaat  | taaaaatctc  | gccgtccatc  | gattttttaa  | taccttaatc  |
| 6961 | aagaataaaa  | caaaaaggtag | gcaacttgta  | tattcaaaat  | ttgacgggtga | ttgcaacttt  |
| 7021 | taactaattt  | caggccattt  | tttgagccgt  | cataactttt  | ttctaaaaag  | ttttcaagaa  |
| 7081 | gtttcattat  | gaaaattcgg  | gttttcagac  | aatltttgagt | ctaataagga  | aataaaaaaa  |
| 7141 | attcgagtac  | accaccttta  | agaaaatttt  | ggattttccgc | tacgctaatc  | cacctttaat  |
| 7201 | caaaaaatatt | tgaagttatt  | caaagttaaa  | gaatttatatt | ttcagagggg  | ccgtagaaga  |
| 7261 | gactcatatt  | tctattttcaa | acgtaagctc  | gatggatatt  | tgaaatgcta  | caaaaacaact |
| 7321 | ggatattttt  | tgttttgtgt  | aagtttttga  | aatacaattc  | gtttgaagat  | ttactctatt  |
| 7381 | ttcagggact  | tttgcacaa   | atgtgggaat  | ttgaccocaga | catcacatat  | aaactgccag  |
| 7441 | cactggaaat  | gtattacaaa  | gagatgtcgg  | aacttgtttg  | tagagaagag  | gtttttgaaa  |
| 7501 | aattttgcacg | agttgcccgc  | atcgcaaaaa  | ctgctgaaga  | tattctgccca | gaggtatgat  |
| 7561 | ttatgagata  | tacagcattt  | cctctaatag  | tattgcatat  | aaacattttca | ctttgaggtt  |
| 7621 | atatcttggt  | ttatttttaa  | aatatcaata  | aatacaaaa   | aatagaaaaa  | tgataaaaaa  |
| 7681 | acattttgtc  | agttgataat  | ttgggtatag  | tattcattca  | taattttgatt | tttttagcga  |
| 7741 | atlttatcgtc | ttgttggtga  | cgtcgaatca  | gctacottga  | gccacaagca  | atgtgctgca  |
| 7801 | cttgttgcga  | gaatgttttt  | tgcccgaccg  | gacagtcctt  | tcagttttctg | ccgggtgagta |
| 7861 | atacaagaat  | gctcatattt  | ttagaatcaa  | tatttgcaag  | gaactttta   | cttacgtacg  |
| 7921 | tcttaagatg  | agcatttttcg | cacatatctt  | acgcgcacga  | gtctcgacac  | gcgaacatcg  |
| 7981 | agcttctgta  | actcgtatca  | atltacaagc  | cgttattaca  | tcagttttta  | atgaatttta  |
| 8041 | agaaaaatcgt | gcaaaaagtag | tgctgagagc  | cattcgcgta  | agatatgggtg | agattttatca |
| 8101 | tttttagacg  | tctagtggat  | atctaacaaa  | actttataca  | tttttatttc  | agaattctct  |
| 8161 | cgtctgataa  | atctattttgt | gtggagaaac  | ttaaattcct  | gttcaacttat | ttcgacaaaa  |
| 8221 | tgtcaatgga  | tccaccggat  | ggtgcccgtca | gttttagact  | tacaaaaatg  | gataaagata  |
| 8281 | cgttcaacga  | agagtggaaa  | gataaaaaat  | tacgttctct  | tcttgaagtt  | gaattctttg  |
| 8341 | atgaaatgct  | tattgaagac  | acagctctct  | gtacacaagt  | tgatttttgcg | aacgaacatc  |
| 8401 | ttgggtggcgg | agtttttaaat | catgggtctg  | ttcaggtagt  | tattttaaagg | aatataagaa  |
| 8461 | tttgaagttt  | tatttttttt  | atgcgaggag  | agatccgttt  | cttgatgtgt  | ccagaatga   |
| 8521 | tggttggaat  | gttggtgtgc  | gagaaaatga  | aaacaactgga | agcgatttca  | attgttgagg  |
| 8581 | cttacgtttt  | cagttcttat  | actggttatg  | gtaagtctag  | actttcaaaa  | aaaactgttc  |
| 8641 | caatatgtca  | atatattttca | ggtcatactc  | taaaatgggc  | agaacttcaa  | ccaaatcatt  |
| 8701 | ctcgtcagaa  | tacaaacgaa  | tttcgagatc  | gttttgagacg | tcttcgggta  | gaaactattg  |



```

12481 tgatattttt ttattaaaaa actgttcgag actacaagaa ggcoctgaaa agcccgcac
12541 gacgactaaa ttcgaaattt cgaattaggt tttaaagatc agaagatcgg cagaaaagta
12601 tctgataaaa atataagaaa tcggaatagg aatgcgatga ggaggtagaa atatggtgaa
12661 gagatacaga agaatagggt aagatcggat gaacttgaag cactttttga gatttttgat
12721 ggtgaagttg gtggatgtag acgtttcatg gaacatctga aaattaagat ttttctaaaa
12781 cacattttct atagaatata atagaatgcc aaatagagaa actagactta cttgaatttc
12841 tttogatttc tgtctttcaa ccttctaact gaaatcaact ttcgacgtgt tctcgggtgt
12901 tcaacaacac catcaacaga acactcagca ccaaattcag catcggaatc atcagaagaa
12961 gactcatcgg aatccaaata gaaattggat ttagtattca tcaattcaaa agaattccat
13021 gatactgtcg attcagcaag ttggactgaa cttgatgggt gactacgaac ccattgaggg
13081 cgctcagggc gaagtcgaga gtatgaggat gcaacgtgga ttgatgatga cgtcaacaat
13141 ctttgggtgt gagatgaaga agtggctgat gcagatgttg acagacggaa tggagatgag
13201 tgaagagcaa gaagacatct gaaaatttga aacgttggtt atgtggacag tactgtaaag
13261 atcttacctt ggatcataac tacttgccct ctgttttctc ttctcttgac ttctacttaa
13321 aagcattttc gtctcgattc tccggttact tgaaaatcca actccagaat ttcagcaca
13381 aagctgctct cccgaaccgt agactgttgc accacgttga ggggttgaca aggatctgaa
13441 atcagatggt taaagcatgg caagtagagc aacaatgtta accaaaattt ctgaaacttt
13501 ttcgaatata gtcaaaaatt gacaataact cagtttcacc tatcatagtt ttggaagtca
13561 accaaaaatt tttgaaattt cataaaaatt ccaaactttc taaaaatttg gaagattgat
13621 atgattgata tgaaagtatt tatatatttt ttaacctggc agacgatact tcaccattaa
13681 agacacacat gtggagaaga attattttac ttttagtaat ccaacgtttg cacttacctt
13741 ggagcatgca agcttttagt cattaaaagc ggaattctag atggagtctt tottgggtgt
13801 gacattgttg aaataaacat tctgtggttc ttgattgatg atgacgtcat agaaccacgc
13861 ccagatgaca atggattacg gtagtcatca gaatcagtag attcattcaa tttctagtc
13921 atttcttctg tttcttgtaa aattaaattt taattaaaga tctaacaaaa atctggcact
13981 tacattaata agataatcaa catattctaa ctcatcctc gtttcattat tttctaattc
14041 tggcttcttc tcatcgaaac gttcgggtggc attgtgtcgt tgcgggcttg accgtttttt
14101 gaatttctga aatgtttttc atgcaatttt tgttcttatt tgtgtgtcat atacagtcaa
14161 aatcaaaaaa tagtacaac taattccgtt tagtaaataa aaaatcgatg taaaatctca
14221 gcaaagccaa gatcttgccg ggtccttata tccaagtttt gttgccattt tatttcagat
14281 attcttttct aaagtcagaa aatttgaatt tagaatcgaa tggacccatt tcttgttttt
14341 ttttgttgca ttttttaact gtaacttttt cgtcagcata tattttcact attaaaacag
14401 aatattcatg acaataattc cacaaaaaaa cgtactttta tatcatagtc gattggttca
14461 gaatttgaac gagaaccttc gacgcgtcga ttgtcagatt ctcgattgat ggacgacgtg
14521 ctgactgaaa atttctggat tgaaaaaata ttcaaatgaa aaaataaatg agaaactcaa
14581 agtctaaaaa atgaatgttg ttaataacga atatttctga tgagaagagg atagagaaaa
14641 aaaacgagtc taataaaatg catgtgatat cctgcataaa aatcccttct ttttctacta
14701 atccttcgct caattcatte aaatagaact ttgatttcta ttagagttga ggttgtttga
14761 acaattttta taaattaaca ataagccata aaacctcgaa acgtaccatc atcattgagt
14821 ttgaaaaagt ggacggtacc gagtcagtc cctctggaac aaatcgttcc agagcactga
14881 aaacgacaac gttctcccca cagaatcgga ttgtctctc ggggaattgt gctcgcacaa
14941 acgatcctga acctgaaaat tttogatttt tgtaagctca atggatttta aactgaaaat
15001 gtagtcaaga agtcaagaaa aactgatgga gttctaaatt cgggtgttagg ttttgaaaag
15061 atcgtcaaac aaacaaatgc ataaaactag gtagggaaca aatagtgaag tagaaaaatg
15121 aaaggcgaca actgccggga gcaagagtac acacaaagaa aaaaagttgc ggaagagcac
15181 agagagcgtc agtccatcag aactgcatag ataaatagat aaagagaaac atgaaacata
15241 agccaccctg ggagagacga caggccagtt ttccgggtgaa gatgagagtg cgagaattag
15301 ataagaaaac ggaaattgtg atgaaacttt ttcaatccaa acttctagaa ttataagaga
15361 cacctaaagt aattagataa gtgttttaag tgatatttta gattcactgt atcatgttta
15421 aaaaagatat ttcaaaaata tatacctgat ataggaggcc tctctgagc accgaattgt
15481 tctcagagctg tttccacgag catccgctca cacattgaca taggccgtcg acagccagga
15541 gttgccacct gaataaataa ttattcaatt taaacctaat ttagtaatgg taactttgta
15601 aatgatggtg gatagctcat ataaaatttg aattggttct aaagttatac aaattttaat
15661 ttcgggtcaaa cttatgaact gtacttttga gttatactat tacaataata ttacccaaat
15721 tattgtatct agatttttgt aatcagtaact aacagatttt aggcaacgtc ctgccagaaa
15781 catgggaata tatttgagca gtttttagta agttgccaca gcttgataaa ggggaattgta
15841 tcaaaatgta cttaataact tctaagcact gacatagtga actacaaaag tccgttattat
15901 acaatgccac tacaaaataa aatattcaga attcgactga aaaatgagaa aaggacaacac
15961 tgaattggac acacgatgtc gtgattttca agaacacaaa aaaaagaaaa agaaatcgaa
16021 aatgttggtt gcctcttttt ctttttctat atgagctaga atctcgaatg catgcctaata
16081 ggagccactc gctctcgttt ctctaagtct cttctcacca gtcttttgtc caaaaattgc
16141 gatgtcgcag gcgtcccgtt tccgcgcctc acggagacac cactatcggc accagatgat

```

|       |             |             |             |             |             |             |
|-------|-------------|-------------|-------------|-------------|-------------|-------------|
| 16201 | cgtgtaaaga  | caccgtcttc  | gttgacttca  | attgctgaaa  | taagaggaat  | tagttttgaa  |
| 16261 | ttggaaatct  | gattaaataa  | aagtcccta   | ttcaatctaa  | ttatttttta  | aacacaaaac  |
| 16321 | ttactattgt  | ttaatgggtg  | tgacgaatta  | gaagaagttg  | attgaaatgc  | gttaacattc  |
| 16381 | cattcaaagc  | tatcttcatt  | ccattctgct  | cgttctttta  | tcctttcact  | cacgtctcga  |
| 16441 | gggatgaagt  | tttcaacaat  | aagaagcctg  | aaaacttata  | ttattctatt  | aaaaaaaatg  |
| 16501 | aacatcaaat  | cctaacgaaa  | gaaaattctt  | ctggggggaa  | aaggagagaa  | ttgtgagaat  |
| 16561 | aaagaacctg  | cgctgtcggg  | atcaaattac  | actatttgaa  | ttcaaattag  | aatacgaaag  |
| 16621 | aaagtgaaag  | aaatgaaaaa  | gagtggagaat | ctattaaatt  | gtaattgaga  | tatcactgaa  |
| 16681 | cttacttcaa  | cttcaattcc  | ttgtcacttt  | ccgcaattgt  | ctgatccaga  | tcttgtcgtt  |
| 16741 | cgtcggaata  | tgctccagac  | acatcacgaa  | tctcattgog  | agcctgaaac  | attcacaaaac |
| 16801 | cttatcttga  | cacctgggtac | atctgaagtc  | aaacctgtct  | caactttatc  | aacatcttct  |
| 16861 | ttaaacttttt | tgtcttttgc  | tccacctctg  | tccttaaata  | cgaaaatggt  | tgtttcaaat  |
| 16921 | ccacagtatc  | ctcttcctgc  | cgttccaatg  | cttcaacat   | ttctctttcc  | cttcgttttt  |
| 16981 | gttcggcaag  | ttctcttctc  | ttcttttcca  | gttgtgcatg  | ttgttctttt  | gtccttgatt  |
| 17041 | ctaacccttc  | atcttcttca  | gatcctaaga  | taagtgcact  | ttgaatgttt  | gctattcttt  |
| 17101 | cagcaaccgg  | tgcttgttct  | atccgttctc  | ttctaaacag  | atcatgcttc  | tcttgaattt  |
| 17161 | ctcttattaa  | tcgatccttt  | tcattgtttt  | tcagtgaatc  | atctttttga  | attgcttcaa  |
| 17221 | tatcatcttc  | aagtttttgc  | cgttcagcat  | cataaaaaact | ttgagttgct  | ccatccctcg  |
| 17281 | atcttgtctt  | tctctgcttc  | agttgctcac  | ggagcatttc  | aatttcttct  | tggaaattctc |
| 17341 | gcagtaaagc  | atccttagga  | tcttcattaa  | ttttcgggtg  | attcttgatg  | tttttagctc  |
| 17401 | gatttgcata  | tcgtaatgta  | ccaagtgtct  | cctcaaaaatt | gtaacttgca  | ggtccaatac  |
| 17461 | aagcaaccat  | aactgtcttt  | gaatttccac  | cgagagaatc  | ttgaagaagt  | cgagtcagtt  |
| 17521 | ttgaatctcg  | ataaggaata  | tgggcagatt  | tcgcatccac  | caatgcactg  | attacatttc  |
| 17581 | caagagccga  | taatgaaaga  | ttgattttcg  | tagcttcttt  | aaatctttcg  | ccagttgctc  |
| 17641 | ctgtttttcga | ttgccgttct  | gaaccagcta  | aatctacaag  | atttagtcga  | ccaactgtaa  |
| 17701 | tatgactttc  | tccgtcttca  | ccaattcggg  | aacattcaac  | agtaatgata  | aagatagcgt  |
| 17761 | gggaacgaga  | cgaatgtctc  | ttcatgttgg  | ttcgccgtgaa | aatttttagta | aaatcaaatc  |
| 17821 | caacggcgac  | cacagaaata  | acttacccta  | cagaccgatg  | cccatttccct | cgaatcatca  |
| 17881 | cttcgtgtat  | ttcacctact  | gtccttggtt  | attttgactg  | aaactttgaa  | atttatagtc  |
| 17941 | gtcttctatt  | tcagaaaact  | atcacttacc  | gttaaactct  | tcacataaac  | tcctccatct  |
| 18001 | ggacgttctt  | taattttctaa | tttcttattc  | gattcggctt  | ctaataaatc  | tcgaagttcc  |
| 18061 | tcctaaagat  | ttcattttttg | taaatcacac  | atcctaaccg  | cttacctgat  | aaatttccaa  |
| 18121 | atagctagct  | ctaactaaat  | actcttgatt  | atgtgatgct  | gccatgtgct  | caaaaaatag  |
| 18181 | gtcaatacac  | ttatagatga  | cacctcgttg  | ttctggatcc  | gatgattttc  | cttccattgt  |
| 18241 | gtgagtcctt  | ccagttccag  | tttgaccata  | tgcaaaaaatc | gtggcattat  | atccgttttag |
| 18301 | aaccgaatca  | actagatctc  | gaaaggtttc  | ttcatataga  | tcogattgtg  | tggaaactata |
| 18361 | aaatatatat  | ttttaaaaaa  | gagaactcat  | aaaatcataa  | acataaaaatt | gtggagaaat  |
| 18421 | aattttgaaa  | aatactaata  | tttctatagc  | agggtgaaaaa | aaagtgatgt  | actcctagaa  |
| 18481 | ataaataaatc | ttactttttc  | tcataaattg  | catcgaatgt  | aaaatccttc  | gatggctcat  |
| 18541 | cttgcctctt  | tggatttttc  | agctcaattt  | gccacggttg  | tggtcgcata  | tgtactatto  |
| 18601 | tgaaaatgat  | cgaattttca  | aaatataaat  | atttcaaaact | ttacttactt  | tgaataatta  |
| 18661 | tttgcaattt  | cttgtgaaga  | taacggtcga  | catctcacia  | ttacctgtaa  | acataaataa  |
| 18721 | atacatttat  | ttgaattttg  | aaatgtataa  | aactggatta  | tgaaattttt  | aagctgggtg  |
| 18781 | tttttgtatg  | agaagtaacg  | aaaaaaagta  | caatttactt  | agagtcttgt  | gatttttctt  |
| 18841 | tcaaattgcaa | aactcaactg  | aatcataaat  | agtgatgctt  | cgaaaagttt  | ttagaggaaa  |
| 18901 | attgtatttt  | tagtaaaaaa  | taataacgt   | tttggactta  | aaaaaaaatt  | atgttaaaac  |
| 18961 | ttgaaaaaatt | acgttttatta | gtgcttatat  | taaaatacgg  | tttcaaatta  | attttaaatt  |
| 19021 | aaaataaactc | accttttggg  | caaaatcaga  | catttttagaa | actagcatgt  | actttattac  |
| 19081 | gttgaatata  | acttatgttg  | gaaaatggaa  | aatttgaaga  | caggtgaatt  | ttagtttttt  |
| 19141 | ttcttttttgc | tacttctaaa  | aaatacttca  | tttatttttac | attttgagaa  | ctaatttttg  |
| 19201 | aacatgtttc  | gaacaaaaaa  | aaagattttg  | aaaaccccaa  | aaaaacttac  | tttgacagtc  |
| 19261 | tcctgttttg  | aagatttttt  | cattattttc  | accatttttt  | gtcactaaat  | atttggctat  |
| 19321 | caatgtaggg  | gtcaaggaaa  | attttggtgc  | attcctgatt  | tagtgagagt  | ggtctggaac  |
| 19381 | ttaagaagat  | tagtttaattg | tggaaaaata  | atcatattgt  | atcgagaaac  | ggaattttga  |
| 19441 | agcaataacc  | gctagagaaa  | gtgactaaaa  | accagaaatt  | gtagtctgtg  | aatttcaata  |
| 19501 | tttttggttt  | tatgtcacat  | ctggacaatc  | ggaaaaatat  | gcatacattt  | gaaattttta  |
| 19561 | gaaatatttt  | gaattaaact  | taaaggaaaa  | aaatgcatta  | aaaagattga  | aaacatcatt  |
| 19621 | gacgttgaaa  | attggagaaa  | atttcttaatt | tctcatcaaa  | atattaaaaa  | atttaaagttc |
| 19681 | ttcaataata  | tgaaaatgtg  | aataaaaatg  | ctaaataaag  | aaaaaaaaca  | gatcctattc  |
| 19741 | attataaaat  | gttcacacaa  | gtgttacatt  | tcgtacaaag  | aagtactaaa  | acggatggac  |
| 19801 | taaagtaata  | ttgtcactcc  | cgaaaagacg  | aggaagaagt  | aatcggaaga  | agatgtcgga  |
| 19861 | agatgagtga  | tagtaaaaaa  | acgaagagac  | gcagatagag  | agtttgagag  | aaggagactt  |

```

19921 ctggaggaat aaaaggtggt ttcaagatgg gggacagaga gggagaggggt taaaagagca
19981 caaaatgtgc ataatatcga tctgtgcgag ttgagagacg cagacaatgt gaagaatgga
20041 gcatatgttt ctagtgaaca ctcagaagta gttgttcatg tgtccgaaac tttggaaaca
20101 tatacatttt aaacttgacg tttttgaatt ataaagggat ggaggtgctt caaaaagtaa
20161 tcatagacat gtgtagatgt taaattaaaa cacaaactaga cataggatga atcagaagct
20221 taccataaca ttgttgatgt atttaaaaat gagaaaaagt aaaattccgg atagtcttct
20281 ttgaaaaaat tcacagagaa gttataatgt ttgatgatat tcaactgattt gtaatacatt
20341 attagtagca tggcttctat gtatatagac tatttttata tcacatacat gaaaaaggggt
20401 taaggcatgc gccagggcct gaaaacgcac ctacctacca ggggagctct agctcttagt
20461 tattaattca agagactttt gaaacttgat tttttgagat tttattcaat gattgggtta
20521 aaaaaaaga ttatttgcaa aaattacaaa ttttaatgtc tatactctga catcgggtta
20581 gaacaatttt agacaggctg caatgaaagc aatgaaataa aatttccttg aaattataat
20641 agagaatcag taaaatgttg cagattatgt gaaaatgcat gcaagaattc gcagaaaatt
20701 cagtgaagca gaaaagtgcg acaggagacc gaagtctaaa aaagtgaatt atgaataaaa
20761 acaaatcatg tgactggata taattgaagg tcttgattcg gaaaagataa ttggagctct
20821 ttgttaggc caggctctag atattttatt gaagcttttc agaattgttc aaatttatca
20881 ggaacagttc tctttgcaat ttctctatgg ctcaactacc agggcttttc ctttttcttc
20941 aaaaagtaga attttaaat ataattttta aatttaataa ccaagcaaaa aatcatatac
21001 tcatcatatc atgtgatcat atcatataat catatagggc togttctttt ttttttcaaa
21061 aaattaaaaa tttactagaa ccaagcatat gacaataaaa tattttgaaat tcaacttaat
21121 gggaaaaaaa caagaaaatt tcattaacat tattgaaaac atogttggca ataggaatgt
21181 agaaaaatcaa atcaaaatca agtgagatta ggaagaatc gaaattaggt agaattggaa
21241 aatctcgatt ttttaagttg gattcttaca cgattttttc gggatatttt tcaattttat
21301 tttgtagtat ttcagcctag acggctgaga attcttttca aaccttccaa tttcaaagag
21361 attcttccat aatttaatat aattttcatt cgatattagc atccattata tacgtatgat
21421 tcccccttta aaatcgattc tctttttcaa acttaagaat tgttgagtca
21481 tcaactgata gtgagcagac accaacaacc tttccgttcc gtttatttta
21541 ttttggaatc taacatatcc aagaaaatta acttgaaatt agaataaatg tttcttgcta
21601 gatttttttg tcataagtat ttcttatttg gattataatt ttcattctga aatcgtagag
21661 agtttttcac tatttttttt tgagttctaa acacttcctt cctcatcgat gatgaagttt
21721 ttgacaaatc aactagtttt ttactcatat ctacatcaa totatgattt tattcaaaaa
21781 cagttaaatt tttttaacga aattaaaatg gtcatcggac cgagcaaaag ctttcagaat
21841 caactgcttc tttaaattct ttaaaattca atcaactttt cgtgtccaaa gtcacaaact
21901 acacctttca aaaaatattt ctacattatt tgcccacatc ttggcacagt tttcttgcca
21961 ttcttcaata ttttcttctc tgcgtttccc acactcttat tttctgactg ttgaactttc
22021 cattgtatag actcaatttt actttcgttt tttcaatttt tttttctgog aagttcgggtg
22081 ttaaacctcc attttgcaat attaaaaatt tcaatattgc oogttttggc ttgaatctat
22141 taaaattatg ctgttttttt ttcagaaaag accaaaacat gccagatgat attccaaaat
22201 tgccacgaca cagaggaaaag aaaaatcagc cgaagggttt gaaaaattta gaaaaatctg
22261 aagttacttt ttttaattct tagacacacc ttggaaacaa caaaaactgc ctgctttacg
22321 gcctcattat aacataactt cagcaattcc agttactctg ataacaggag tagocacgtt
22381 ggcaatggga attgctcttt atttcggaca taatggatgt gagtttttag agtttattat
22441 cccaaaaaca aaaatatcaa ttactctttc ctggtataaa gtaagaaaaa gctaaagaaa
22501 acaaatttct tgtcaaaatt ttacattgta aaccgatagc aacaaaaaac aagtgtcata
22561 aaaactgtaa gaaaatcgat aatttttgta caatttcaca aagctaaaaa atatttttta
22621 ttttacgtg ttagtaccg aatgttctgc acttgagcct tactattagt tacacaaaat
22681 ggatcaattt tgagcaattt gttgtgaatc tgacaataat tagtctatt gatatagctt
22741 taggccactc attcgtgttc gtaattttcg ttttccttga acttgtaaag gtacagtttt
22801 tgaaaacagg gatgtagtcc aagtagtcaa atattgattc ttgtagcatt agaacaagag
22861 attgtgtaca cggattgtgc tctttcaaat ggaacacaag cttcacgaat tatgagaact
22921 gaaatgggaa atcaaacatt taaatgtgca tatacaatta ctttgaatga cgattatact
22981 gtaagttgag ttttaatttt taaatcatca aagaaaacat atgtatattt ttgogaagga
23041 aattttggat ctggtcttag gatgaaacga cttgttaaca ttttgattaa agagccotta
23101 gttggaagtg agtgtatctg gtaaaaacac aattcgaaaa tatttaacca aatatgtata
23161 aagcctaggt tgaacctgct ctgcagttcc taatttttca cattattttt cttcaaaaata
23221 ttactatgat atttcaaagc ccgggggtacc atcttaaaat catcatttgc aagtatcaca
23281 attaatgttc aacattacag ggcgaaagtga agttttatta cggctctttc aagttctatc
23341 aaaacaatcg attatacttc aactcacgaa acgatcaaca gctacgtgga aaagttactg
23401 aaactgacgg atgtgatcca ttagaatatg tggatgttaa ttggaactaaa ttgcccattg
23461 cgccgtgtgg gaaagtggct gattcaatgt ttaacggtgc atttcaattg attgottaat
23521 ttcagtattg caacattttt cattttttat aatacatcta acttcaaaat ttgttttttt
23581 ttcagatacc ttcgaattat tttatatcaa tgataaagcc tcaaacggcg taacacgggt

```

```

23641 tccatggaca actcgtggag tactcgggtgc aactgaaatg aaaagaaaaat tcagaaatcc
23701 gattcgcagcg gaaaaccaga cattatgtga tgtgttttgcg gttgaaatga aataagaaaa
23761 aaaataaatt aaactccatc ttttagggaa caatgcctcc gccatcatgg agatatccga
23821 tctgtcaatt gggactaaac agtattgata cagatgttgg cattgggttc gagaacattg
23881 attttatggg ttggatgaag gttgcagctc ttccaaaatt cagaaaaactg tatagaatac
23941 tgaatcgaca agttgatatg ttcagtaatg gattacctaa aggacaatat cagttgacca
24001 ttaattacag tatgtttatg ttaatgttga atttatgtat ttatgcaaaa aatttactgc
24061 aaaagttcac aataattcca cccaaacctg cttaaatatg gagatgcaag ttttttgttt
24121 cagataaaca gtggctccaa aaaaccaatc tttgttataa aaocctcaca aaatttctcg
24181 atattttctt ttctttcata ttgatttggg aaacttttga gaaaaaaggg aaatttagaa aattctttca
24241 agcgaaaattg tcaaaatttt tcaaaaccaa atttgatttt ccagatttat tttttgtcga
24301 cttgacaata gtaaagaaaa aacaagttga atttttctat atgaattctt atagctgaac
24361 atttttgatc aatttgaaaa taatcaatag acaatttttc ccatactac tgattttcag
24421 actatccagt ggatatgtat tcgggcgaca agtacttctg tatagccaat gaaaaactggg
24481 ttggaccagg gaatctgttt ctaccagtaa tctatttggg tgttggaaac ttcttacttc
24541 tcttaactat ttctttcata ttgatttggg taaaacagag actgtcgagg gttcatccaa
24601 catgaattgg aaaaactaat tgaaaataga cggatgaact tcaaatttgt ttacaagagt
24661 tgaagtctca aaataagctg gtagcatgta ttgtacggga acagatttgt atactttgct
24721 ttgtaataaa aataaaatgt tattatatta gtctgtaatt ttatgtatag ttcaatttaa
24781 ttgaaatata taataccccc ttcagtttat caattaaagc tccaactatc attcgttggg
24841 tgagattaat tgtcgcagtga gggcatctga aatgtaaatt taaaattaca aaataaataa
24901 ttgtaaagtc tatcagatat acaaaatgat catttaatta aggaggaaaa acaaaacatt
24961 aattttaaaaa atttatcaaa aaacaaaaaa aaacggtcaa atattttttc aaacaaacaa
25021 agtaaagcta atttctatta aagttgatct aaatactggg tgtgtaggca tactatagtt
25081 gatttcaacg ggaagaaggc caaatcagca agtgtacatt gtgttctgaa aaattgaaat
25141 tcaacagttg aatataagta gaaactctac ctatttgctaa catttattgc aattcttctg
25201 tgttttgaac aatatcgaga tcgctccatc catcggataa ttccgtatga tttgatgaca
25261 tctcatctac agcttccaat tctccaatta tctgatcctt cagtttccagc ttaatatcaa
25321 acgatttttg aatttctctg atatttgcct cataaactct cgagatttca gatttaactt
25381 gttgaatctc tttaaattgc tctgaaatct tcttctcaag aactcttttc tgataatata
25441 gctcagtgat ttctcgagtt ctttctttca taatcatttc agtattcatt tgttcttctt
25501 taacattttc ttcaagtttt tcaatctttc gtacataatc acagaaatga tcaaccacct
25561 gcaaccattg cgggtcattt cgcatagttt tgagtcctcc gggttgttca agaattgcga
25621 caagactttc tgtctgttca agtttcagtt tttctaagtt ttgctcaagt ggggaattctc
25681 tgacctctcg gtttgcctct tcggaaggat cagaaaattt ttccagaatg aacgtcaaaa
25741 tttcttgttc caaatgggga tatgttctac tagtcccttg actagaagtc tcaactagaga
25801 ttttaagtg cagtttctga acattacgct ccaatttcgg agaatttcca gatcagtg
25861 tcacatgact ctttaatttg acaatctcat cattcttctc aaagatttga ttttccagtg
25921 cagagacctt cacttgacaa tcttttgcct tccatgacaa ttcatctgcc aacatcttga
25981 tcatgttctc attcgaatca atcgtttttt tccatctttt tccatctttt tatcatgtct
26041 tgattgtaac attttgcttt gaaatttccac gtttccacaga attcaaagca attttcaaag
26101 aattgttgaa attttcgagc atcaaaacttt caaacacctt ggatttatca ttcttggcaa
26161 cattccgatc attatttagt tcagttgata tagaattggg atcatcaact gagaatactg
26221 ttaaattctg gtgttgaagt tccaaaaatt tctccatgac tttttcaata ctatcttctt
26281 gataaacatt gaaggattct tgaagcattt taatctcttc attctttttc ttgagttcgg
26341 aaaacagaat attctttctt ttttggaaatt ccaactgtcac taagagataa tcatctttgt
26401 tttccaaaga tgtaaccagc tgcgatggat tcaaaacttt caagtttttc gtaagatagt
26461 caacttcgtc ttcaagtttt tgaatgatct gtttgttgtc ttccatcttt tttgttataa
26521 gtataggatc cttgaaataa agagaaaaac tgactatgat ottgtcaata gtttccagca
26581 agtggaatat gtcataataa tctccatcat ttaatacttt taatttgtcc aaaagttgat
26641 ccatcaactt ctcttgtttc ggtgtttcat caccaattaa tattccaccg taaccgttaa
26701 ccgcatattg cgacaaatca taaaattggt tttgaagatc ctcatattta gttttgagaa
26761 ttgagaaatc ttcatctctc attatcaaca acttcttcaa tcttctagct tctgtctctg
26821 cttttatatt gaaaatttct aattgcattt cagctagtat tatctcatct tcttgttcat
26881 cgattaattt ctcaacttct tottttaatt ttttgatatt ttgatctttt tctgccatag
26941 cattttgaaa atcttctagt gttgcaaaact gatattcagt ggatgtactc gtggattctt
27001 gagtcgaaat ttcagtggtc ttgtttatct gttttacttc ggagcttctg actattttct
27061 ctggagtcca aaacttgtct acttccaaaa aatgtgtttt tttgcttctg aaaaacatat
27121 actaagtaac atctttaaga tattcagggt gggttccagt aagcaagcat aattctgaaa
27181 actttgaatt atcaatctga attcttctgc aogaaacaac tattgtattt aaaaatgct
27241 ataaaaatta cagcttttga aaccaatgaa tctttttttc acacttttcc agtttatcaa
27301 cacttcaact ccattctctt ccaccgctgc

```

```

27361 ttaaaaaatc aagtttctgt tgttcaggtg aaggctgaga tgcgtggaac gacatagttc
27421 tgaaaatagt taatttaa atgtacagaaa aatcttttct agaaagtaaa aaaaatcagt
27481 aaaaacaagt actaagagaa attgaataaa ccaatcacaa taatgacttc ttaacaagct
27541 gaaaaataat gcaatagcaa agaaaaacga gtagtttcgg taactccata gtacattatt
27601 tegtatttgg gatcatcata tcatttattg atgaggatat tatgagttaa ttctaataac
27661 ccgagagtaa aggcaaaaa tagcatggag tgaaaaaacg gatcaagcaa agaaatcgtg
27721 ttaactttta taacatctag ttgacactgt cagaccaaaa acttaataaa attttcaactt
27781 gtacataaca gctagctgaa actgtaattt aattttatat tccctcggtc aattctagct
27841 aaatttagcg ttctgagcta agccttcatt tcaaaattaa caaaaaaaat gcaatgaaat
27901 tttcacttgt acataacagc tagctgaaac tgtaatttaa ttttatattc tctcggtcaa
27961 ttctagctaa attagcgatt ctgagctaag acttcatttc aaaaaataca aaaaaaatga
28021 attgaaattt tcacttgtag ataacagcta gctgaaactg caatttaatt ttatatctct
28081 tcagtcattt tcagctaaat tagcaatttt gagctaagtg ttgttggttc ttaaaacaat
28141 gcaaattttg atggtttttc gtgttcagtg aacaaacaaa caaacaacaaa aattctggta
28201 aataaccaca agctgaaact gtgagataat tttttagtga ccattgagtg actgctcata
28261 gacagtggct tggaaattaag actagaatga ttatctctca tgataacata ttatacagag
28321 aagttgggaa gaatgtaggt cattgtaaag cgacagacag gtcgcattga tcaaagagaa
28381 tataagtcga actctttcgt ttggtaactt gagggccaat gttatttgct attagggaaa
28441 attaacattt aaggagcaaa ggattgcaaa caaatgcga taagatatat gattatagta
28501 ttttatcttt tgtaagtgtt gccataattt cagtaacgaa aaaaaataca aggcaatttt
28561 agatgttagg aaaatcgaat ttgtctgact agccaacgaa tgttctcaat tgaagttatt
28621 gttctttttt aagatgtttt catacaaatt agtcagtttt cgaagcttca gcaacactta
28681 tccgaattga gcaatttcaa aactattttt tgtaaaataa aatacatctc cgaaaattta
28741 catcgagttc ccaacaatac tgtatggata gaaaatacct accaatactg cacatgaaac
28801 gctctgaaaa taatcgga aaatgaga acctttttaa tataaaatga gcacaataag
28861 taataactaac tttattgaga aagaacataa ttgttatgag aatagttttt aaatgagggtg
28921 agaaacagaa tatccctgag aataagttaa gatacttgaa aatttggtga atagtaataa
28981 gtaaaatggt ttacacattag tataaacaat gacagagtca cgcaaaagta cgggaaacat
29041 atgaagttta taatacagtg cagtacagaa aaggtaacaa gtttacaaga atacaattgt
29101 tttttaaaaa taattttttg ttgaaggctt aaggtaatac gattaaagag ctactttctt
29161 ccaatacga gttgaattta aaatttaaaa ggaaaaaagg aaaaaaatta aaaagcatat
29221 gaaaaatcgg ggcgcathtt tagtgcaaaa aattagatgg catttatttt atcccatcca
29281 tctgaatctt cactgtgtgt ggatttattg tgcgcattct gatcgatcat tgtatcatca
29341 gcttctcctt cttgattgat aagaagacct tgcagttttt ccgaaagtcc cgaaatcttc
29401 aaatccttct ctctcaatgc atcatgcac ttctgaattt cagcgatcgc ttcgctattt
29461 tgaataagtt ccatcagaca ctcaattttg ctatcttttt ccatgatttc tcttttatga
29521 ttgcaatct gttcttcttt tgattcacat tctctctttg aattggctga aataaaagaa
29581 aatgcttaca gatgtgtgta aaacccttag aaaactttca caagcttacc tgtcaatact
29641 tcaaattgcc ccaataagtt gtgcttccac tcttcagttc gaagtttaag atcttcaact
29701 gatgtattaa gcgtggcttt ttctgctga gtgtttgcaa gttgcatctc taacgccatg
29761 acggtcgagt tatgttgatc caaaatatga ctg

```